UNIVERSIDAD AUTÓNOMA DE MADRID
FACULTAD DE CIENCIAS
DEPARTAMENTO DE BIOLOGÍA

EVOLUCIÓN ESPACIO-TEMPORAL DEL GENERO *HYPERICUM* (HYPERICACEAE): LA ESTRATEGIA DEL CORREDOR DE FONDO

MEMORIA DE TESIS DOCTORAL

ANDREA SÁNCHEZ MESEGUER

MADRID, 2013

MEMORIA PRESENTADA POR LA LICENCIADA
ANDREA SÁNCHEZ MESEGUER
PARA OPTAR AL TÍTULO DE DOCTOR EN CIENCIAS

DIRECTOR DE LA TESIS:
ISABEL SANMARTÍN BASTIDA

Este trabajo ha sido realizado en el Real Jardín Botánico de Madrid, CSIC, bajo la dirección de Isabel Sanmartín, científica titular del Departamento de Conservación y Biodiversidad. Además, parte de los resultados fueron obtenidos en la Universidad de Gotemburgo (Suecia) bajo la dirección de Bernard Pfeil, y en el Field Museum of Natural History (Chicago, USA) bajo la dirección de Richard Ree. La realización de la Tesis has sido posible gracias a una Ayuda predoctoral de Formación de Personal Universitario (FPU, AP-2007-01698) concedida a Andrea Sánchez por el Ministerio de Ciencia e Innovación (ahora Ministerio de Economía y Competitividad) y al proyecto titulado "Reconstrucción del origen de la Rand Flora Afro-Mediterránea con datación filogenética e inferencia biogeográfica" (proyecto CGL2009-843 13322-C03-01/BOS) concedido a I. Sanmartín por la misma entidad.

Cover image: *Hypericum androsaemum* L. Ariedge, Pyrenees, France.

A mi madre
A mi padre

"Lo mejor para la tristeza – contestó Merlín, empezando a soplar y resoplar –es aprender algo. Es lo único que no falla nunca. Puedes envejecer y sentir toda tu anatomía temblorosa; puedes permanecer durante horas por la noche escuchando el desorden de tus venas; puedes echar de menos a tu único amor; puedes ver al mundo a tu alrededor devastado por los locos perversos; o saber que tu honor es pisoteado por las cloacas de inteligencias inferiores;

Entonces solo hay una cosa posible, aprender.

Aprender por qué se mueve el mundo y lo que hace que se mueva. Es lo único que la inteligencia no puede agotar, ni alienar, que nunca la torturará, que nunca le inspirará miedo y desconfianza y que nunca soñará con lamentar.

Aprender es lo que te conviene.

Mira la cantidad de cosas que puedes aprender: la ciencia pura, la única pureza que existe. Entonces puedes aprender astronomía en el espacio de una vida, historia natural en tres, literatura en seis. Y entonces, después de haber agotado un millón de vidas en biología y medicina y teología y economía y geografía e historia, pues, entonces puedes empezar a hacer una rueda de carreta con la madera apropiada…"

Terence White, The Once and Future King, Putnam's Sons, Nueva York.

TABLE OF CONTENTS

TABLE OF CONTENTS

AGRADECIMIENTOS

AGRADECIMIENTOS

Ante todo, me gustaría agradecer a Isabel Sanmartín el haberme dado la oportunidad de poder realizar la tesis doctoral bajo su dirección. A Isabel se lo tengo que agradecer todo: primero el confiar en mi, segundo ser tan generosa, tercero tener siempre tiempo y ganas de enseñarme algo, cuarto por transmitirme su motivación y sus inquietudes científicas, y quinto y último, por haber hecho que esta tesis se convierta en una aventura emocionante. Si Isabel te mira y te pone ojos pícaros, tiembla, porque ya sabes que va a proponerte hacer algo que conlleva una carga indecente de trabajo, pero a lo que seguramente no te vas a poder negar porque va a merecer la pena. Ni soñando podría haberme imaginado tener una directora asi.

A mi tutora académica, Virginia Valcárcel, por facilitarme la vida con los trámites de la tesis.

Quisiera también hacer extensivo este agradecimiento al resto de miembros del grupo que, durante los primero tiempos (Juan Jose Aldasoro, Cajsa Lisa Anderson, Marisa Alarcón, Sven Buerky) y también durante los últimos (Mario Mairal, Lisa Pokorny) me han proporcionado una ayuda sin lugar a dudas clave para acabar esta tesis.
También a Jorge Lobo, el inicio de nuestra colaboración ha marcado un antes y un después en la trayectoria de esta tesis.

También quisiera expresar mi gratitud a los equipos de investigación que me han acogido en sus laboratorios durante las estancias breves, por la gran experiencia que ha supuesto poder trabajar con ellos y lo mucho que he aprendido.

En primer lugar el grupo de investigación de Bernard Pfeil en la Universidad de Gotemburgo (Suecia): Thomas Marcussen, Bengt Oxelman, Yann Bertrand and Filipe De Sousa.

Y por otro lado el grupo de investigación de Richard Ree en Field Museum of Natural History (EEUU): Matthew Nelsen and Deren Eaton.

Muchas gracias a todas las personas que han contribuido a la realización de este trabajo enviándome muestras *Hypericum* o recolectando material para mi: Juan Jose Aldasoro, Julian Molero, Andreas Hilpold, Pablo Vargas, Jose Ruiz, Alejandro Gonzalez, Cristina Tauleigne, Miguel Serrano, Bengt Oxelman, Mario Mairal, Borja Mila.

Gracias a los diversos herbarios que me han proporcionado mucho del material que se ha incluido en esta tesis: MO, A, BOZ, AAH, GB, W, AAU, BCN, BC, UPS, S, pero en especial al herbario del Jardín Botánico (MA) y a Mauricio Velayos y Concha Baranda que me han ayudado enormemente con el préstamo de los especímenes.

Gracias especialmente a Isabel Marques que tuvo paciencia y me enseñó todo lo que hay que saber en el laboratorio, y también a los técnicos de laboratorio del Real Jardín Botánico, Fátima Durán, Gemma Andreu, Gillermo Sanjuanbenito y Emilio Cano que han hecho posible este trabajo al colaborar en la extracción de ADN y su amplificación. También gracias a Yolanda Ruiz por su buen hacer en el uso del microscopio electrónico.

También me gustaría dar las gracias al resto del personal del Real Jardín Botánico por

haber proporcionado todo el soporte necesario para realizar este trabajo, a los miembros de la biblioteca, al personal de administración, a los vigilantes, a los trabajadores del Jardín. Gracias en especial a Maria Aguilar, Ricarda Riina, Javier Fuertes, Angélica Bello, Jose Luis Blanco, Isabel Marques, Jano González, Juan Carlos Zamora, Elena Amat, Arantxa Martínez, Raquel Pino, Isabel Liberal, Irene Villa, Luis Valente, Lifang, Melisa Sarmiento, Vladimir, … y un largo etc por crear un excelente ambiente de trabajo. Además, muchos de ellos han aportado nuevas ideas, críticas constructivas, y tantos buenos consejos.

Gracias a quien corresponda por la beca FPU que me fue concedida y sus ayudas para estancias breves, a los proyectos de investigación que han financiado mi trabajo, y a las becas de viaje que me han permitido asistir a varios congresos.

Gracias a los que formaron parte de mi carrera científica antes de esta tesis (Antonio Rosas, Samuel García Vargas, Almudena Estalrich, Antonio García Tabernero, Marcus Bastir, Elena Kranioti, Fátima Franco y Pablo Manzano) y a los que forman parte de mi nueva vida científica y personal en Montpellier (Emmanuelle Jousselin, Arnelle Cour d'Acier, Sophie Veran, Maria Razzauti, Marie Pages, Fabien Condamine, y otro largo etc…).

Por último, y de manera muy especial, gracias a las personas que han estado muy cerca. Guillermo, Sara, María, Irene, Ana, Celia, Guzmán, Borja, Andreas, Marieta, mi tía Chelo, Loli, Jordi, Joan, A Martin, que ha llegado tarde a esta tesis pero a la vez ha llegado a tiempo, le tengo que agradecer todo, y en especial que me mantuviese con vida mientras mi tobillo agonizaba. A mis padres, mi hermana, sobre todo.

PRESENTACION EN CASTELLANO

GENERAL INTRODUCTION

PRESENTACIÓN EN CASTELLANO

Con el fin de reconstruir la historia evolutiva del género *Hypericum* esta tesis incluye conceptos y herramientas provenientes de diferentes disciplinas, como la Sistemática, la Filogenética, la Biogeografía, la Paleontología y la Ecología. El objetivo principal del trabajo que hemos realizado es investigar los patrones de distribución y riqueza de especies, poniendo un énfasis especial en el efecto que tienen el clima y el cambio geológico sobre los tres procesos que modelan la diversidad: la especiación, la extinción y la dispersión. A lo largo del manuscrito mostramos las limitaciones que tienen los enfoques que utilizan únicamente la información presente para reconstruir el pasado, ya que la diversidad actual puede estar sesgada y no ser representativa de la fracción de diversidad extinta. Igualmente destacamos el papel que pueden desempeñar los fósiles para solventar este problema.

Cualquier intento de generalización se basa en la extrapolación de los resultados obtenidos en grupos concretos de organismos. Para el cometido de esta tesis hemos seleccionado el género *Hypericum* (Hypericaceae), ya que cumple varios requisitos como organismo modelo. *Hypericum* es un género grande (c. 500 especies), con distribución cosmopolita (esta presente en los cinco continentes) y con un registro fósil que se remonta al Eoceno superior. Por otro lado, es el género más rico en especies del clado tropical "clusioide" (Orden Malpighiales), al que pertenece, y el único miembro que se dispersó y fue capaz de diversificar en el Holártico durante los periodos de inestabilidad climática que caracterizaron el Cenozoico hasta el presente.

Existen numerosos estudios sistemáticos del género *Hypericum*, y a día de hoy, se le considera uno de los géneros mejor conocidos desde un punto de vista taxonómico y morfológico. Norman Robson estudió el género en profundidad y publicó una serie de 12 capítulos monográficos donde describe 488 especies clasificadas en 36 secciones morfológicas (Robson, 1977, 2012). Por el contrario, el conocimiento que existe de las relaciones evolutivas del grupo es mucho más limitado, quizás en parte por la dificultad de trabajar con un grupo tan grande y de amplia distribución. En la última década han surgido estudios filogenéticos basados en marcadores moleculares, aunque la mayoría de ellos tienen limitaciones por el reducido número de especies estudiadas y porque se basan en un sólo marcador molecular, el interespaciador nuclear ribosomal ITS (Park y Kim, 2004; Crocket et

al., 2004; Hennan et al., 2008; Pilepić et al., 2011). Procesos evolutivos como hibridación, introgresión o duplicación génica pueden generar discordancia entre el árbol de genes y el árbol de especies. En este sentido es importante construir una hipótesis evolutiva sólida, que incluya un muestreo representativo de especies y que esté basada en varias regiones moleculares que a su vez abarquen distintos genomas. Las filogenias así producidas servirán de marco para testar hipótesis biogeográficas, ecológicas y de evolución de caracteres.

En el **Capítulo 1** de esta Tesis analizamos las relaciones evolutivas de *Hypericum* y grupos cercanos en base a marcadores nucleares (ITS) y plastidiales (*trnL-trnF, psbA-trnH, trnS-trnG*), usando inferencia bayesiana y de máxima verosimilitud. El muestro de este trabajo incluye un 40% de las especies del género y un 90% de las secciones morfológicas. Analizamos la congruencia que existe entre las filogenias basadas en distintos marcadores moleculares y proponemos nuevas medidas para mejorar la estima de longitudes de rama en inferencia Bayesiana en aquellos casos en los que existe una alta heterogeneidad en la tasa media de variación molecular entre marcadores. Los resultados de este trabajo muestran que el género *Hypericum* no es monofilético, sino que incluye otros géneros de la tribu Hypericeae (*Triadenum*, *Thornea*). Encontramos un bajo nivel de congruencia entre las secciones morfológicas tradicionales y los resultados filogenéticos moleculares. La filogenia presenta una estructuración claramente geográfica, donde la mayoría de las especies se agrupa en dos grandes clados: Nuevo Mundo y Viejo Mundo. El clado del Nuevo Mundo incluye principalmente especies americanas de las secciones *Brathys*, *Trigynobrathys* y *Myriandra*, mientras que del Viejo Mundo tiene su centro de distribución en el Paleártico e incluye especies de todas las demás secciones del género. Estos dos grandes linajes aparecen relacionados con poco soporte con un clado pobre en especies del oeste del Paleártico, secciones *Elodes* y *Adenotrias*. Por último, la filogenia producida nos sirve de base para inferir el tiempo de origen y la historia biogeográfica del grupo, así como la evolución de algunos caracteres morfológicos diagnósticos. Así estimamos que el ancestro de *Hypericum* fue un arbusto sin glándulas negras, originado en el oeste del Paleártico al final del Eoceno (ca. 35 Ma), probablemente a partir de ancestros africanos. De allí dispersó al Neártico en el Oligoceno, y más tarde, a finales del Mioceno, a todos los continentes del hemisferio sur. Lo más probable es que esta última dispersión se realizase a través de las montañas tropicales que se elevaron durante el Neógeno. Sin embargo, esta reconstrucción podría estar sesgada, ya que presenta discrepancias con el registro fósil del grupo (ver más abajo). La

reconstrucción de caracteres ancestrales sugiere que muchos de los caracteres diagnósticos utilizados en la taxonomía son el resultado de fenómenos de convergencia o paralelismo evolutivos.

En los últimos años un número creciente de estudios señala que es necesario considerar evidencias moleculares independientes para producir hipótesis filogenéticas robustas (Doyle, 1992). En este contexto, y como complemento al primer capítulo, en el **Capítulo 2** evaluamos la utilidad que tiene estudiar genes nucleares de copia simple para reconstruir las relaciones evolutivas de *Hypericum*. Para este estudio testamos diferentes regiones de ADN y encontramos que, entre todas ellas, en *Hypericum* solo PHYC y EMB6572 amplifican con éxito para diferentes grupos taxonómicos. Estos nuevos marcadores son prometedores, ya que producen filogenias consistentes, con niveles de resolución y soporte comparables o superiores a otros marcadores nucleares y plastidiales clásicos.

El **Capítulo 3** de esta tesis representa el primer estudio paleobiológico del género *Hypericum*. En este capítulo realizamos una revisión de la literatura existente sobre fósiles de *Hypericum* para mostrar que el fósil más antiguo, *H. antiquum*, data del Eoceno superior y está localizado en Siberia. En esta misma área han aparecido otros fósiles que datan del Oligoceno inferior, aunque han sido asignados a diferentes especies extintas distintas a las encontradas del Eoceno. A partir del Mioceno existen numerosos restos en yacimientos de varios continentes. En este trabajo también discutimos los principales caracteres diagnósticos de las semillas y del polen de *Hypericum,* ya que son los fósiles más abundantes en el registro; y en base a estos caracteres revisamos la asignación del fósil más antiguo, *H. antiquum*. La edad y distribución de los yacimientos con restos fósiles de *Hypericum* nos permite aproximar la paleodistribución del grupo, mostrando que *Hypericum* se originó en el este del Paleártico, durante el Eoceno, y que posiblemente formaba parte bien del bosque boreotropical o de su sucesor, el bosque mixto-mesofítico, que cubrió el Holártico desde el Paleoceno hasta el Mioceno.

En el **Capítulo 4** hemos desarrollado un nuevo enfoque integrativo basado en la combinación del registro fósil con inferencia biogeográfica y modelos de nicho ecológico. Este enfoque nos permitió inferir el nicho climático de los ancestros, testar eventos de evolución de nicho, así como identificar áreas de distribución ancestrales y potenciales rutas de dispersión que en el pasado fueron adecuadas para las tolerancias climáticas de *Hypericum.* Contrariamente a los resultados obtenidos en el Capítulo 1, la inclusión en el análisis de fósiles y preferencias

ecológicas de linajes ancestrales nos permitió reconstruir un escenario más realista de la historia biogeográfica de *Hypericum*, congruente con el registro fósil. En concreto, los resultados sugieren que durante el Eoceno los ancestros de *Hypericum* estaban distribuidos por todo el Paleártico formando parte del bosque boreotropical y que, posiblemente, utilizaron el corredor de Beringia para dispersar al Nuevo Mundo. Además, inferimos que el nicho global del género se ha mantenido estable desde el Eoceno hasta el presente, aunque algunos clados se han especializado en distintas dimensiones del nicho en vez de explorar nuevos espacios ecológicos. Los resultados de este trabajo tienen implicaciones más allá de *Hypericum,* mostrando que las preferencias ecológicas de los organismos pueden preservarse durante periodos de tiempo de decenas de millones de años, lo que representa una evidencia a favor de la hipótesis de conservación de nicho, por la cual las plantas tienden a conservar sus preferencias ecológicas en el tiempo evolutivo (Donoghue, 2008).

Por último en el **Capítulo 5** nos preguntamos cómo se ha originado la extraordinaria diversidad que presenta *Hypericum* en el presente y abordamos este estudio a través de una combinación de técnicas de mapeo estocástico de caracteres, modelos ecológicos de nicho y modelos estocásticos de diversificación. Desde antiguo, los biólogos se han sentido atraídos por estos linajes hiperdiversos y han tratado asociar tales radiaciones a la aparición de innovaciones evolutivas que pudiesen facilitar la adaptación de estos grupos a su ambiente. Aunque probar causalidad es difícil, establecer estas correlaciones es crucial ya que permite investigar las causas que produjeron la acumulación de linajes y descubrir porque unos grupos diversifican y otros no. Con casi 500 especies *Hypericum* es uno de los 100 géneros más grandes de las angiospermas, y el más diverso del clado clusioide al que pertenece. Además, *Hypericum* es un rara avis dentro del clado clusioide ya que es el único linaje distribuido en regiones templadas del hemisferio norte, y junto con la familia Podostemaces, el único que ha evolucionado un hábito herbáceo en algunas especies. Estudios recientes muestran que los ancestros del clado clusioide eran probablemente plantas leñosas de bosque tropical cerrado en Gondwana (Davis et al., 2005). Este resultado sugiere que, pese a que el nicho de *Hypericum* se mantuvo estable desde el Eoceno hasta el presente, en algún momento previo de su historia evolutiva los linajes ancestrales experimentaron una evolución en sus preferencias ecológicas y de su morfología que les permitió adaptarse a los nuevos ambientes templadas, fríos y estacionales, que aparecieron en el Holártico durante el Cenozoico. En este sentido se han propuesto dos hipótesis: La primera sugiere que el éxito evolutivo de

Hypericum se basa en la aparición de innovaciones evolutivas tales como el cambio de nicho o de hábito que produjeron aumentos rápidos en la tasa de diversificación del grupo ("key-innovation diversification" KID hipótesis). La segunda hipótesis, "time-to-speciation effect" (TSE), sugiere que la edad del grupo y la lenta acumulación de linajes a través del tiempo pueden explicar por si mismas las diferencias de riqueza de especies entre *Hypericum* y sus linajes hermanos. Los resultados de este trabajo muestran que a lo largo de su historia evolutiva *Hypericum* experimentó varios cambios de preferencias ecológicas, algunos coincidentes con importantes eventos climáticos al final de Eoceno. Sin embargo, ni los cambios de nicho ni la aparición de formas herbáceas durante el Mioceno estuvieron acompañado por aumentos en las tasas de diversificación. De hecho la riqueza de especies de *Hypericum* es la esperada para su edad y la diversidad del grupo puede explicarse por la acumulación constante de linajes a lo largo del tiempo. Como un "corredor de fondo", *Hypericum* muestra una capacidad excepcional de adaptarse y lidiar con el estrés (resiliencia), posiblemente gracias a una gran plasticidad genética ya presente en los linajes ancestrales. Esta plasticidad podría haberse mantenido hasta el presente, lo que permite explicar el éxito evolutivo de *Hypericum*.

GENERAL INTRODUCTION

For centuries, biologists and naturalists have noticed that the number of species on Earth is not randomly distributed both at local and regional scale in marine and terrestrial environments, but tends to concentrate in certain geographical areas or regions (Humboldt, 1820; Darwin, 1845; Wallace, 1852; Candolle of Geneva, 1805). 200 years later, the mechanisms and processes underlying the geographic variation in species richness are still subject to intense debate. This area of research is indeed one of the central themes in Biology. Diversity emerges as the interaction of ecological and evolutionary processes acting through time (Ricklefs, 2006), but the factors regulating these processes are intensely debated (Benton, 2009; Ezard et al., 2011). Two main mechanistic models have been proposed: The "Red Queen" model is based on Darwin's view of evolution, a balance between different biotic pressures, and attributes the regulation of species diversity to intrinsic factors such as biotic interactions (van Valen, 1973). Conversely, the "Court Jester" model postulates that changes in speciation, extinction and dispersal rates occur in response to extrinsic abiotic factors such as climatic changes (Barnosky, 2001). These two models probably represent two extremes of a continuum and in reality biotic and abiotic drivers of diversification may prevail at different temporal and spatial scales (Jablonski, 2008; Ezard et al., 2011). At large and long temporal scales the regulation of diversity is generally attributed to extrinsic abiotic factors such as climate or geological change (Lomolino et al., 2005; Benton, 2009; Wiens et al., 2010). Because the signature of biotic interactions might get saturated and is difficult to observe, patterns of biodiversity appear driven by the physical environment at this scale.

Climate is probably one of the most important factors of the physical template, as there is only a specific set of climatic conditions under which species can persist. This is defined as part of the species fundamental niche, the set of conditions that allows a species to maintain viable populations (Hutchinson, 1957; Wiens & Graham, 2005). In general, species tend to conserve their ancestral niche preferences over evolutionary time (Crisp et al., 2009), but in some exceptional cases niche evolution occurs, usually by developing key innovations ("adaptive breakthrough" or "key adaptation") that allow an organism to rapidly diversify and the invasion/adaptation of new environmental conditions different to the ancestral. The niche is an important concept in biology, given that the ability of an organism to evolve

ecological tolerances under scenarios of climate change – the interplay between niche conservatism and niche evolution – will determine the capability of the species to adapt to new environments. Several genetic and environmental factors determine this plasticity, but ultimately it will depend on the speed of the change and on the geographic distribution of the organism. Indeed, there is a reciprocal effect between niche dynamics and biogeographic processes. Niche preferences directly influence the ability of organisms to disperse and colonize new regions (Wiens et al., 2010), while the availability of migration corridors will mediate into the balance between dispersal and evolutionary change (Donoghue, 2008). Under the general principle that it is easier to move than to adapt, if climatic corridors exist, species will tend to disperse to more favourable areas and preserve their climatic niche preferences (geographic sorting; Herrera 1992; Ackerly 2004). Conversely, if corridors are missing, then relevant adaptations to the new environment will presumably evolve (Donoghue, 2008).

Together, niche dynamics and biogeographic evolution play a major role in the generation of diversity. Niche evolution, also known as "ecological innovation", has been related to events of rapid diversification or adaptive radiation, such as the colonization of new mountain niches by lowland organisms (Hughes & Eastwood, 2006) or the transition from marine to freshwater habitats. Niche evolution can be associated with geographic speciation - the colonization of new geographic regions (invasion of new areas after developing the relevant adaptive traits followed by speciation in allopatry), but it is more often related to sympatric speciation via ecological displacement or selective evolution (i.e., selection of two different optima in the same geographic region, Snichzler et al., 2012). Niche conservatism might also promote cladogenesis through dispersal to new geographic regions with similar climatic tolerances (within the climate range of the organism that disperses) and subsequent allopatric speciation. In this understanding, niche dynamics could be considered an additional evolutionary process shaping species distributions and providing causal explanations for diversity patterns (Wiens et al., 2010).

Despite the close association between niche dynamics and biogeographic processes (Wiens & Donoghue, 2004), it is surprising that the ecological preferences of organisms have been generally overlooked in biogeographic analyses. In part, this is explained by the scientific tradition of studying Ecology and Biogeography as unrelated disciplines, separated by different methodologies and underlying assumptions. However, to a certain extent, this is also

a consequence of the popularization of the Neutral Theory in the field of Ecology, built on the idea that the presence of species in a particular area is a stochastic process and that ecological differences are not important in maintaining diversity (Hubbell, 2001). On a more practical level it is related to methodological difficulties to explicitly incorporate "ecological connectivity" into biogeographic scenarios.

In more recent years, Ecology and Biogeography have closed positions through the use of Ecological Niche Modelling (ENM) techniques applied to phylogenetics and the concept of "ecological vicariance" (Sanmartín, 2012). Vicariance has traditionally been defined as the result of allopatric speciation in response to the appearance of a geographic barrier. However, evolutionary ecologists such as Wiens (2004) redefine vicariance as the outcome of any environmental change that causes a division in a species geographic range. ENMs use the association between distribution data (species occurrences) and environmental variables (e.g., temperature, precipitation) to predict the range within which a species could occur (Kozak & Wiens 2010; Sanmartín, 2012). Assuming that niches are preserved over time ("niche conservatism"), and given information about past climates, one can project back the ecological niche for different points in time to reconstruct past species distribution and ancestral dispersal corridors across regions that are now uninhabitable (Yesson and Culham 2006; Weaver et al. 2006). Furthermore, the assumption of "niche conservatism" can be avoided through the combination of niche models with the fossil record of extinct lineages (e.g., Maguire and Stigall 2009), so these models can be used to trace patterns of niche conservatism and evolution over longer time scales (Stigall 2012).

I. Cenozoic climate change and its effect on Holarctic vegetation

The history of flowering plants could be traced back to the Cretaceous, when most of the present diversity was originated in a rapid burst, grounding of the decline of other land plant forms (Magallon & Sanderson, 2001; Crane and Lidgard, 1989). By the beginning of the Cenozoic, angiosperms dominated terrestrial plant communities (Lidgard and Crane, 1988), on a greenhouse world characterized by tropical temperatures at northern latitudes (Zachos et al., 2001) and several land corridors connected today isolated landmasses (Tiffney, 1985a,b). For example, the North Atlantic land bridge connected North America and Europe (Tiffney, 1985ab) and the Turgait strait divided the Palearctic into an eastern and a western side (Sanmartín et al., 2001). At this time, a uniform vegetation belt, the boreotropical forest with

no analogous in the present, extended across the northern hemisphere through a narrower Atlantic Ocean and a warmer Beringia (Wolfe, 1975; Tiffney, 1985a,b). The Paleocene-Eocene Thermal Maximum (PEET) at 55.8 Ma was a significant warming event on the entire Earth, which was accompanied by an increase in plant biodiversity that far exceeded Holocene levels (Willis & MacDonald 2011). From this time onwards, the Cenozoic has been by a period of long-term climate cooling, with a gradual decrease in global temperatures punctuated by some warm intervals (Zachos et al., 2008). At the end of the Eocene, a dramatic drop in Earth temperatures, the Terminal Eocene Event (TEE), led to increases in aridity and seasonality, and promoted the selection of cool-adapted boreotropical elements, and the ultimate expansion of grasslands and a temperate deciduous vegetation, the "mixed-mesophytic forest" (Tiffney, 1985b; Fig. 1), over the previous extension of the boreotropical belt. The cooling trend continued in the Miocene. After a brief warm episode (the Mid Miocnene Climate Optimum (MCCO), a new drastic decline of temperatures at 11 Ma, the Late Miocene Cooling (LCM) event, led to the expansion of a boreal coniferous forest across the northern regions of Eurasia and North America (Schneck et al., 2012), whereas temperate plants became restricted to climatic refugia in the south (Xiang et al, 2000). Around 3.5 Ma ago, temperatures went up again, especially in the high latitudes, in what is known as the Mid-Pliocene Warm (MPW) interval (Zachos et al., 2008). This was presumably accompanied by large-scale range shifts in many plants, with cool-temperate deciduous forests expanding northward at the expense of the boreal forest. In the Mediterranean region, however, a new, more xerophyte vegetation assemblage was developed, i.e., the Mediterranean-type biome. Since the Quaternary (the last 2 million years), the Earth's climate has alternated between glacial and interglacial conditions (an "icehouse world"), leading to expansions and contractions in geographic ranges, which were more rapid and frequent than in previous periods (Willis & MacDonald 2011).

Figure 1. Reconstruction of the mixed-mesophytic forest in eastern North America, American Museum of Natural History, New York. Photo courtesy: Isabel Sanmartín.

In the face of the Cenozoic climatic changes, the vegetation of the Northern Hemisphere responded in three different ways (Willis & MacDonald 2011): some groups became extinct or restricted to climatic refuges (Wolf, 1975; Wenn, 1999; Tiffney, 1985; Xiang et al 2000; Donoghue & Smith 2004); others migrated to the south tracking their preferred habitat (niche conservatism), while others became adapted to the new habitats and found in them an opportunity to diversify (niche evolution; Fig. 2). The latter appears to be a rare event, according to phylogenetic evidence. Only seldom tropical plants succeeded to make the transition to temperate habitats, i.e., cool and highly seasonal environments (Judd et al., 1994), since this required the modification of complex physiological systems (Donoghue 2008). Indeed, large-scale ecological conservatism seems to be prevalent in plants, suggesting that ecological tolerances have remained stable over long periods of time (Qian & Ricklefs 2004; Crisp et al., 2009; Donoghue 2008). This has been argued to explain the global Latitudinal Diversity Gradient, the observed decrease in species richness from the tropics to the pole in numerous plant and animal groups.

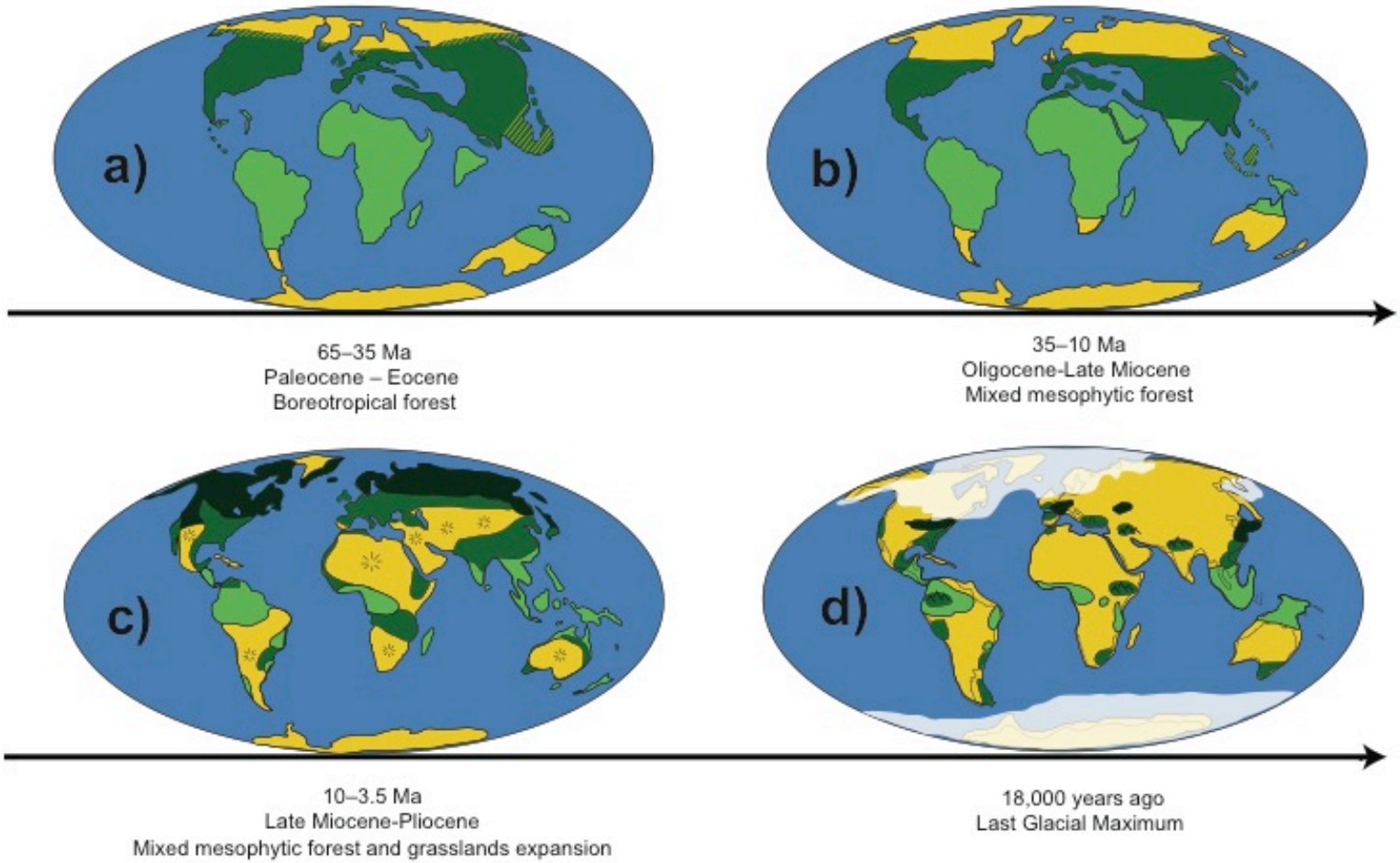


Fig. 2. Scheme representing changes in vegetation composition in response to major shifts in Cenozoic climate. **a)** During the warm Early Eocene, a "boreotropical" forest, composed by hardwood deciduous and broad-leaved taxa, extended across North America and Eurasia. **b)** After the TEE event, this forest was replaced by a temperate, deciduous mixed-mesophytic forest, whereas evergreen boreotropical elements migrated southwards or became extinct. **c)** Global climate cooling starting in the Late Miocene led to the expansion of a boreal coniferous forest across northern Eurasia and North America, displacing the temperate forest. **d)** During the Quaternary, ice sheets covered large regions of Eurasia and North America and a tundra belt extended in the north; temperate plants became restricted to climatic refugia in the south (see Text for a more detailed explanation).

Most of these plant lineages originated in the Late Cretaceous-Early Tertiary tropical conditions and presumably preserved these climatic tolerances over evolutionary time, while only a few of them were able to develop the relevant adaptations to the more seasonal climate of temperate latitudes (Wiens & Donoghue 2004; Donoghue 2008).

One such lineage that managed the transition from tropical to temperate climates (Donoghue, 2008) is the plant genus *Hypericum* (Hypericaceae), source of the active compound known as hypericine, which is used as a painkiller or antidepressant in western traditional medicine. Although *Hypericum* has a cosmopolitan distribution, most species are distributed in the Northern Hemisphere, while some lineages reach the tropical mountain regions of South America, Africa, and South East Asia. (Robson, 1981; Fig. 3). With nearly 500 species (496), *Hypericum* is considered one of 100 largest angiosperm genera in the world (Carine & Christenhusz, 2010). It is also the most diverse within the group it belongs to, the "clusioid" clade from order Malpighiales, which includes 5 families (Hypericaceae, Bonnetiaceae, Calophyllaceae, Clusiaceae s.s. and Podostemaceae), in 94 genera and circa 1900 species (Ruhfel et al, 2011; Wurdack & Davis, 2009; Gustaffson, 2002).

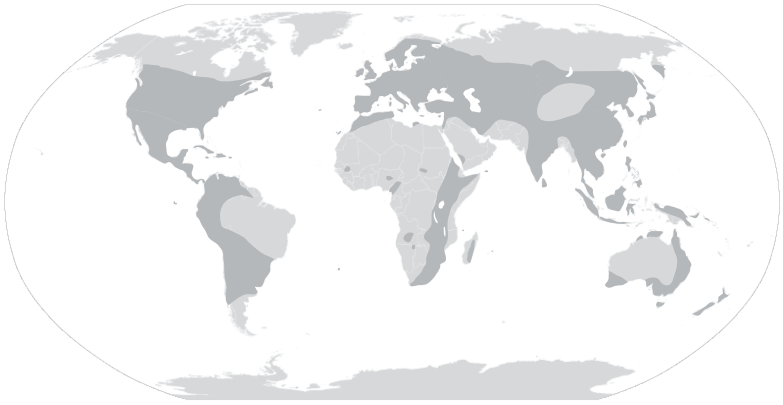


Figure 3. Present distribution of *Hypericum* species. Map showing the current distribution of *Hypericum* species (modified from Robson, 1977).

Most species in the clusioid clade, as well as in other malpighial families (e.g., Malpighiaceae, Davis et al., 2002), are tropical woody species inhabiting the rainforests of South America and Southeast Asia (Davis et al, 2005). Davis and collaborators (2005) suggested that members of this clade evolved from tropical ancestors that inhabited close-

canopy rain forests in Gondwana (Davis et al., 2005; Ruhfel, 2011). However, new biogeographic inferences and the fossil record suggest that some of these clusioid lineages were already present in the Holarctic in the Late Cretaceous and Early Tertiary, probably favoured by the tropical conditions existing at that time (Ruhfel, 2011). There are fossil remains of clusioid clades from the Late Cretaceous in North America (*Palaeoclusia*, a stem relative or crown member of the clusioid clade, Crepet & Nixon, XX; Davis et al., 2005; Stevens, 2007; Rufhel 2011), the early Tertiary of Siberia (*Hypericum antiquum*, Hypericaceae, Arbuzova, 2005), and Middle Eocene-Early Oligocene of Southern Europe (*Calophyllum* pollen, Callophylaceae, in Spain; Cavagmetto & Anadón, 1995). During the course of the Cenozoic, most of these clusioid lineages presumably went extinct from these latitudes, i.e., all extant descendants are tropical clades, with the exception of *Hypericum*, which seems to have survived and diversified in the temperate regions of the Northern Hemisphere, where it presents today its highest species diversity (Robson, 1981). Another interesting aspect of *Hypericum* is its wide range of habit forms: together with Podostemaceae, *Hypericum* is the only clusioid lineage that exhibits the herbaceous habit and this change in life form between woody to herbaceous might have been key to its large diversification (Smith & Beaulieu, 2009).

Understanding the mechanisms generating and maintaining diversity necessarily has to be based on the extrapolation of results inferred from a few selected groups of organisms. Several of the aforementioned aspects make *Hypericum* a good model system: i) *Taxonomic knowledge*: *Hypericum* is one of the few large plant genera with a complete taxonomic treatment (Robson, 1977; 2012), which greatly facilitates the selection of species for biogeographic and phylogenetic analyses, something necessary in a group of this size. ii) *Ancient age*: Wurdack & Davis (2009) suggested that all families of Malpighiales diverged in a rapid burst around the Mid Cretaceous (110), with divergence between *Hypericum* and its closest relative *Vismia* dated at 65 Ma; the oldest fossil record of the genus is from the Late Eocene (Arbuzova, 2005). This implies that *Hypericum* or its stem lineages would have lived through all the major climate changes of the Cenozoic, including the PEET, TEE, LCM, and MPW events. iii) *Cosmopolitan distribution*: the genus is present in every continent with the exception of Antarctica and covers a wide range of habitats (Robson, 1981), which allows comparison of diversity patterns across distinct geographic regions and climates. iv) *Niche and morphological evolution*: the invasion of the Holarctic by *Hypericum* ancestors and the

appearance of the herbaceous habit are potential key innovations ("adaptive breakthroughs") that could drove the diversification of the genus.

Altogether, these characteristics make *Hypericum* an ideal model group to study the interplay between niche dynamics, geographic evolution, and diversification within a single lineage. The ultimate aim is to understand the role of climate change on species origination and distribution, which might help us infer how they will react to future changes. Given the role of human-induced climate change in the current biodiversity crisis, this is today is a key area of research in Biology. Below, we describe in more detail the genus *Hypericum* and which was the knowledge on several different areas before the start of this thesis, and our objectives in relation to this.

II. The genus *Hypericum*

The plant genus *Hypericum* L. comprises at present 497 species of trees, shrubs and herbs (Robson 2012; Figura 4). The name derives from the greak *hypér* = over ad *eik´ōn, -ónos* m. = image; i.e. over everything you can imagine, in part because of its reputation as a medical plant but also associated with the use of some species, placed on religious sites, as protection against evil spirits. At present, some species like *H. perforatum* (Saint Jonh's wort) are economically important for its medical properties. Specifically, the active compounds, hypericine and pseudohypericine, produced in the dark glands, have proven antiviral and anticancer properties (Matzk et al., 2001). *H. perforatum* is also one of the most popular anti-depressive treatments in natural medicine, and its historical use is documented before 2400 years ago. In addition, other members of the genus like *H. androsaemum* o *H. calycinum*, are frequently used in horticultural practices to construct hedges and fences.

Species of *Hypericum* are characterized for having black and red glands containing hypericine and pseudohypericine, as well as amber squizogenous cavities filled with resins and essential oils. Dark and pale glands appear both in reproductive and vegetative organs and in some cases forming vesicles. Stems have a variable number of lines or ribs (0)2-4(6) between nodes. Leaves opposite, sometimes whorled and without stipules. Generally, leaves are entire and could present auricles or being fimbriae at the base.

Flowers hermaphrodite. Perianth 3–5 merous, free to partially fussed. Petals yellow, but some species present white or red corollas. Stamens numerous and generally grouped in fascicles with filaments free to partially united. Fascicles free to partially united forming a continuous

Figure 4. Morphologic diversity in different *Hypericum* species

ring. Some species are unique to present 3 or 5 squamiform staminoids (vestigial fascicles) alternate between the fertile stamen fascicles. Carpels vary in number from 3 to 5 with the styles free or united. Fruit a dehiscent capsule but some species present fleshy indehiscent drupes. Seeds numerous cylindrical to ellipsoidal, inferior to 1 mm (Robson, 1981; Ramos-Nuñez, 1993)

2.1 Taxonomy and classification

Since its description, *Hypericum* has been the subject of intense taxonomic work, but especially during the last decades (Choisy, 1821; Spach, 1836a, 1986b; Jaubert and Spach, 1842; Engler, 1925; Keller, 1925; Kimura, 1951; Robson, 1977, 1981, 1985, 1987,1990, 1996, 2001, 2002, 2003, 2006, 2010a, 2010b). The taxonomic adscription of the genus has long been discussed. Early classifications show important differences in rank, relationships

and nomenclature: Choisy (1821) did the first treatment of the genus and defined subfamily Hypericineae, with tribe Hypericeae representing *Hypericum* in its current sense. Spach (1836a, 1836b) considered Hypericaceae Juss. a family and divided it into two tribes, the tribe Hypericeae Choisy y Desmostemoneae Spach. From tribe Hypericeae he excluded *Elodes* Adans and *Adenotrias* Jaub. & Spach, based on their particular morphology: the presence of vestigial fascicles in these species is shared with other Hypericaceae tribes but not with other *Hypericum* members. Engler (1925) and Keller (1925) defined subfamily Hypericoideae Engl., with three tribes; Hypericeae, Vismieae Choisy y Cratoxyleae Benth. Tribe Hypericeae comprising two genera, *Hypericum* L. and *Ascyrum* L, differing in having 4-5 merous perianth respectively. Kimura (1951) followed Spach and excluded section *Elodes* from subfamily Hypericoideae, dividing the latter into four tribes; Hypericeae, Androsaemeae Kimura, Sarothreae Y. Kimura (including genera *Myriandra* Spach, *Brathydium* Spach and *Sarothra* L.), and tribe Ascyreae Y. Kimura with genus *Ascyrum* L. Norman Robson did the most exhaustive taxonomic revision of *Hypericum* to date, describing 488 species and defining the main diagnostic morphological characters of the genus in a series of 12 publications (1977, 1981, 1985, 1987, 1990, 1996, 2001, 2002, 2006, 2010a,b y 2012). He returned to Engler (1925)'s classification, and considered Hypericoideae Engl. a subfamily inside Guttiferae, including tribes Hypericeae, Vismieae, and Cratoxyleae. Robson also divided the genus into 36 morphological sections, including the species-rich American *Trigynobrathys* and *Brathys*, the Asian *Ascyrea* and the temperate *Hypericum*.

The current classification, based on DNA phylogenetic evidence (APG III, 2009; Stevens,

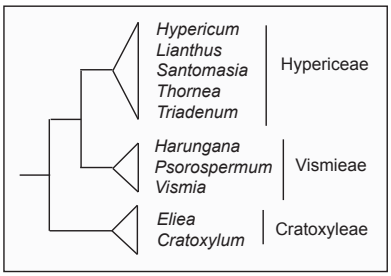


Figure 5. Schematic representation of phylogenetic relationships among the genera of family Hypericaceae, showing division into tribes.

2005; Ruhfel et al., 2011; Nürk et al., 2013) recognizes *Hypericum* as a genus inside family Hypericaceae, which includes three tribes: Hypericeae (*Hypericum, Triadenum* Raf.*, Thornea* Breedlove & McClintock*, Santomasia* N. Robson and *Lianthus* N Robson), Vismieae (*Vismia* Vand.*, Harungana* Lamarck and *Psorospermum* Spach.), and Cratoxyleae (*Cratoxylon* Blume and *Eliea* Cambess: Fig. 5). However, the circumscription of Hypericeae and

relationships among the tribes remains controversial. Recent phylogenetic work suggests that *Hypericum* is not monophyletic, with other genera, *Santomasia*, *Triadenum* and *Thornea*, nested within (Nürk & Blattner, 2010; Ruhfel, 2011).

2.2 Evolutionary relationships

Robson was the first to propose an evolutionary scenario for *Hypericum* based on hypothesized evolutionary trends for the main morphological characters (Robson, 1977; Figura 6). Based on Robson's study, Nürk and Blattner (2010) carried out the first cladistic morphological analysis of the genus and discussed the morphological evolution for all species of *Hypericum* described in Robson's monograph. They confirmed the monophyly of some sections (*Campylosporus*, *Androsaemum, Roscyna, Sampsonia, Olympia, Origanifolia, Coridium, Myriandra* and *Adenotrias*), but their phylogeny did not recover many other relationships proposed by Robson (1977).

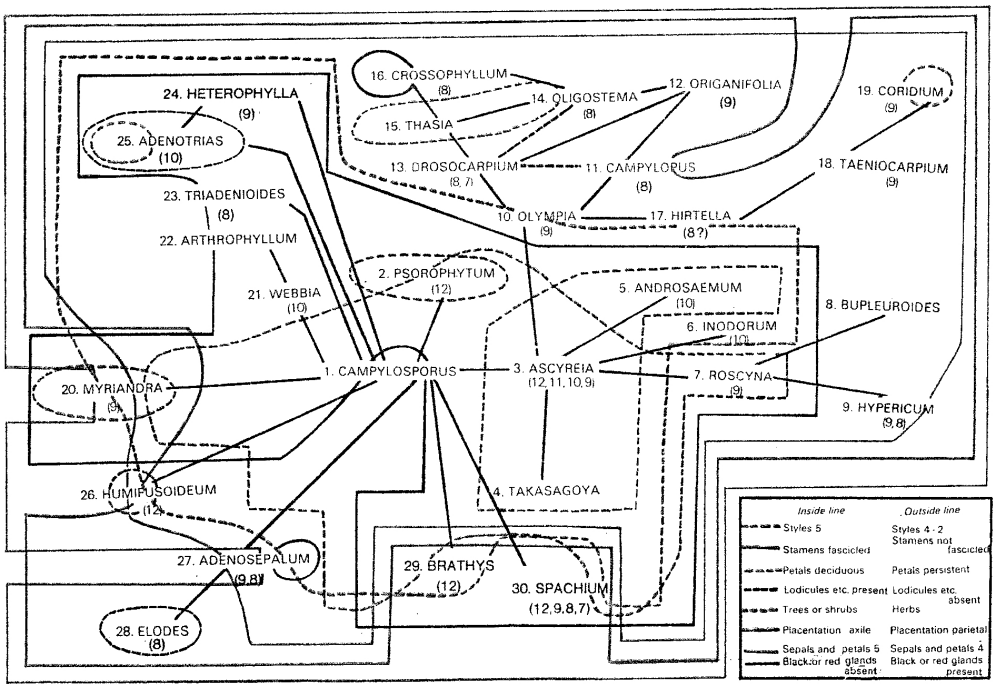


Figure 6. Schematic representation of evolutionary relationships in *Hypericum* as proposed by Robson (1977), showing the evolution of some diagnostic morphological traits. Numbers in brackets indicate chromosome counts. Following Robson, all current species were originated from the section *Campylosporus* in Africa.

They also found discrepancies with Robson's evolutionary trends. For example, Robson (1981) suggested that herbaceous and shrubby species in *Hypericum* evolved from tree like species, while Nürk & Blattner (2010) inferred a shrubby ancestor. On leave venation, Robson suggested that parallel venation is the plesiomorphic state, while Nürk & Blattner (2010) inferred a pinnate ancestor. However, Nürk & Blattner (2010)'s results need to be carefully interpreted, since basal relationships suffered from lack of support and resolution.

Phylogenetic studies of *Hypericum* based on molecular markers did not appear until the last decade, partly because of the difficulty to study such a large and widespread genus. Thus, the first studies were rather limited in species numbers and taxonomic coverage. Park & Kim (2004) analysed evolutionary relations in 36 species from Korea and Japan. They suggested that the section *Hypericum* is polyphyletic, although their results were poorly supported. Hennan (2008) added 3 recently described species from New Zealand to Park & Kim (2004)'s sampling. The work conducted by Crocket et al. (2004) is more representative and includes a biggest geographic coverage, but still sampled only 50 species. Crocket and collaborators recovered all the study species grouped in two clades, one including species from the American section *Myriandra*, and the other grouping Old World species from several sections. Similar results were obtained by Pilepić et al. (2011), who included 33 species (17 new accessions), although these authors recovered *Myriandra* as non-monophyletic. Ruhfel and collaborators (2011) published a multigene phylogeny for the clusioid clade, including 20 species of *Hypericum*. They found that the genus is not monophyletic, comprising species from the other Hypericeae genera: *Triadenum*, *Santomasia* and *Thornea.* Recently, Nürk et al. (2013) published the first comprehensive phylogeny for the genus, sampling more than 200 species (40% of the genus diversity). They confirmed the inclusion of *Triadenum* within *Hypericum,* but recovered *Thornea* as the sister group of *Hypericum*. Excepting Ruhfel et al. (2011), all phylogenetic studies published to date has been based on a single molecular marker, the nuclear ribosomal intergenic spacer (ITS). The recognition of the difference between gene trees and species trees (Doyle, 1992) has discouraged the use of a single molecular marker in phylogenetic reconstruction (Hillis, 1995). In fact, this represents one of the biggest challenges in phylogenetic reconstruction, as the discrepancy between different types of evidences is frequently the result of biological processes like hybridization, duplication or incomplete lineage sorting.

In **Chapter 1** of this doctoral thesis, I present the first comprehensive phylogeny for *Hypericum* based on the nuclear ribosomal marker ITS and three non-coding chloroplast markers, the interspacers *trnL-trnF*, *psbA-trnH*, and *trnS-trnG*. Taxonomic sampling was greatly increased with respect to previous studies, covering 40% of the species of the genus (186 of the 496 species) and 33 of the 36 described morphological sections. Phylogenetic results published to date do not fully support the traditional sectional classification with many sections recovered para- or polyphyletic. In many cases, discrepancies between taxonomy and phylogeny could be due to homoplasy in morphological characters, resulting from convergent evolution, producing errors in the diagnosis of plesiomorphic states. Nürk & Blattner (2010, 2012)'s results suggest that this is the case for *Hypericum*. To test this hypothesis, I conducted ancestral state reconstructions for some diagnostic traits in *Hypericum*, such as the presence of dark glands, the habit form or the shape of the corolla. This was done in a hierarchical Bayesian framework, which allows simultaneous estimation of phylogenetic relationships and ancestral states, while integrating out the uncertainty in tree topology, branch lengths, and ancestral states ("mapping uncertainty", Ronquist 2004).

The analysis of different types of evidence, such as morphology or the nuclear and plastid genomes might allow us to identify important biological phenomena like convergent evolution and hybridization. To this purpose, low or single copy nuclear genes represent a vast and unexplored source of phylogenetic information, being especially appropriate to solve interspecific relationships and to study hybrid speciation (Sang, 2002). However, up to date, their uses in systematics have been limited, probably because of the methodological and theoretical challenges to clarify the complex evolutionary dynamics of nuclear gene families, including the difficulty to differentiate between paralogues (product of a duplication event) and orthologous (product of speciation) gene copies. Lack of universal primers is another difficulty to work with these markers. In **Chapter 2** I evaluate the usefulness of single-copy markers in phylogenetic reconstruction of *Hypericum*. The objective of this chapter it to identify variable regions and design specific primers that could be used in future phylogenetic studies to increase resolution and branch support values at basal and interspecific relationships within *Hypericum.*

2.3 Biogeography

Phylogenetic and biogeographic data from extant organisms can help us infer when and where a lineage originated, while ecological evidence informs us on extant climatic tolerances. The fossil record can add to this information by revealing us new geographic areas and/or ecological conditions in which the lineage's ancestors lived, but which are now unobservable because of extinction. Rarely, this information has been combined into a single approach to understand the factors fuelling the engines of a lineage's evolution. The phylogenetic relationships presented in **Chapter 1** constituted the evolutionary framework for conducting subsequent analyses. By combining the phylogeny together with fossil evidence, molecular dating, ancestral state reconstruction techniques and niche models, we aimed at disentangling the spatio-temporal framework for *Hypericum* diversification (Figure 7).

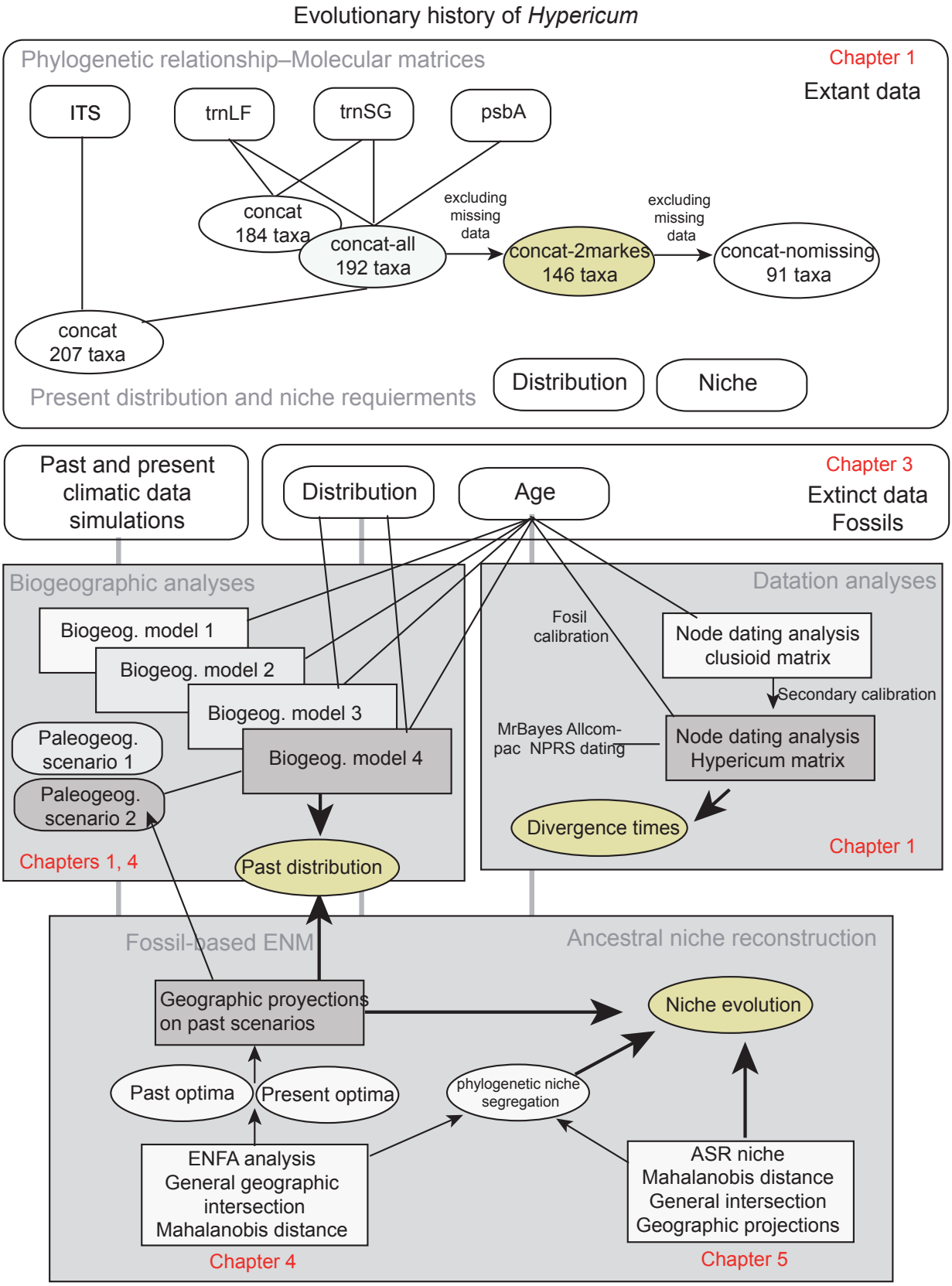


Figure 7. Flowchart of the work

At present, *Hypericum* exhibits a nearly cosmopolitan distribution. The largest diversity is found in temperate regions of the Northern Hemisphere, but it also occurs in tropical and subtropical mountains of the Southern Hemisphere. *Hypericum* is only absent in the poles, arid deserts, and low-altitude tropical areas (Fig 3). However, this species richness is not evenly distributed among sections and geographic areas (Fig. 8). The most species-rich region is the Central and South America, with 104 species. The Mediterranean region has 95 species, Asia (82) and the Irano-Turanian region (80). Nonetheless, it is worth mentioning that most South American species are grouped within two morphological sections, while all the species richness in the Mediterranean region belongs to twenty different sections (Fig. 8). The genus also presents large biogeographic disjunctions. In Africa, related species are distributed in the margins of the continent, separated by the Sahara desert and the Central African tropical lowlands, in what is known as the Rand Flora pattern (Sanmartín et al., 2010; Fig. 3). In North America, members of the section *Trigynobrathys* have its closer relatives in Africa (*H. lalandii*) and Asia (*H. japonicum*).

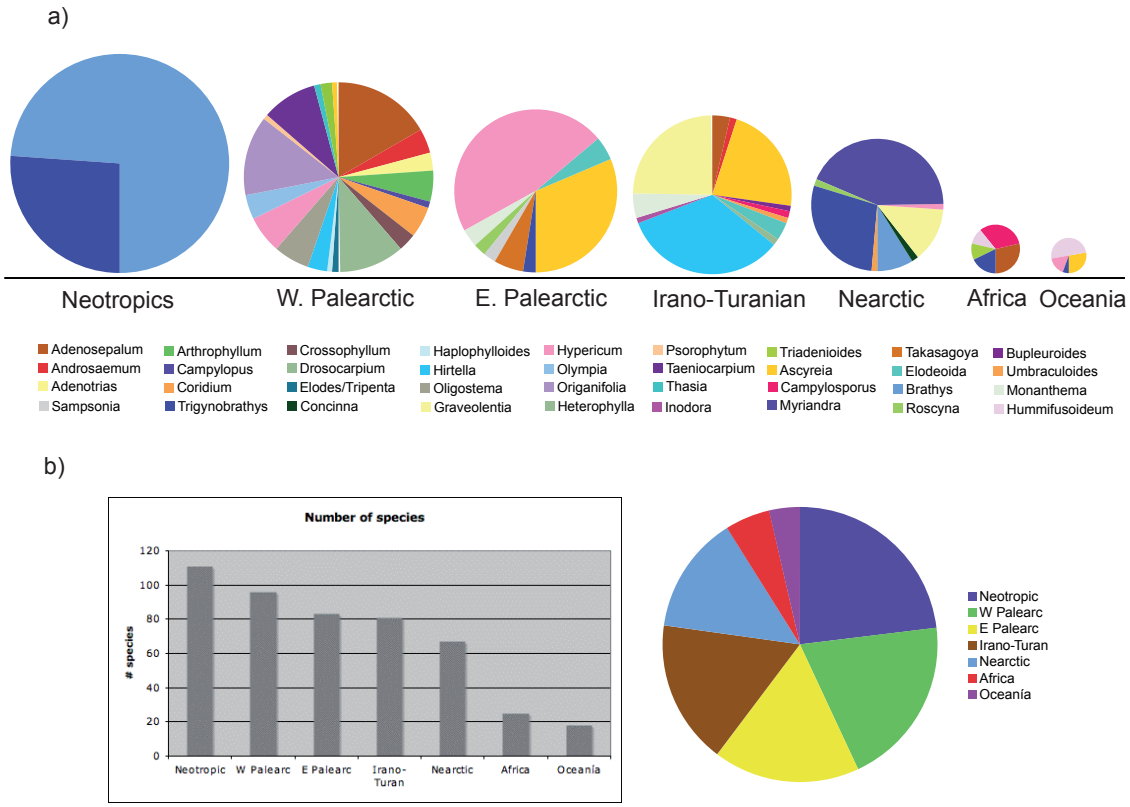


Figure 8. a) Pie charts representing the proportion of species of every morphological sections distributed in the biogeographic regions considered in this study. Sizes of the pie charts are scaled according with the richest area (Neotropics) and proportional to them. Pie chart colours correspond with morphological sections. **b)** Total number of species per biogeographic region.

Hypericum species also exhibit a great heterogeneity in habit forms: some species are dominant trees in African mountains, others are herbs in Central Asian grasslands or aquatic adapted ryzomatous herbs in Western Europe, although the majority of species are shrubs. In the Holarctic, species are present across a wide range of habitats, including open habitats (e.g., *H. perforatum*), riverbanks, damn and wet areas (*H. humifusum*, *H. tretrapterum*), rocky shelters (*H. aegypticum*), or mountain slopes (*H. ericoides*, *H. montanum*). The Southern Hemisphere is characterized by an extraordinary diversity concentrated in tropical and subtropical mountains, where *Hypericum* is abundant in the subalpine belt. More precisely, the Andean Paramo hosts nearly 88 species, 50 species occur in the Himalayan range, and 20 in the Eastern African mountains.

So far, several hypotheses have been proposed to explain the wide but uneven biogeographic distribution of *Hypericum*. Robson (1981) thought that the group was very old and probably originated in Africa before the connection between Gondwanan landmasses was broken in the Mesozoic (Robson, 1981). However, this hypothesis disagrees with current fossil evidence (Arbuzova, 2005) and with recent molecular dating analyses (Davis et al., 2005; Ruhfel, 2011), all of which suggest an early Cenozoic origin for the genus. Nürk & Blattner (2010) proposed a Palearctic origin for *Hypericum* based on the basal position of Mediterranean clades in their morphology-based phylogeny. However, none of these hypotheses have been tested within a proper statistical framework, partly due to the lack of a robust phylogeny, but also due to the fact that the fossil record of *Hypericum* has been until we started this work rather overlooked (Robson 1981).

To overcome these limitations, in **Chapter 3**, we conducted a survey of the literature of *Hypericum* fossils and discuss the main diagnostic characters of seeds and pollen (the most abundant remains). We also identified the oldest *Hypericum* fossil described to date, *H. antiquum* from the Late Eocene of west Siberia (Arbuzova 2005), and revaluate its assignation. Finally, in this chapter, we discuss the paleogeography of *Hypericum* based on fossil occurrences.

This new information was used in **Chapter 1** to reconstruct the time of divergence and biogeographic origin of major *Hypericum* lineages, using a statistical approach that allows simultaneous estimation of these parameters and phylogenetic relationships. The biogeographic model used in this chapter is based in present distribution of the species to

infer ancestral areas in the past. It suggested a Western Palearctic origin for *Hypericum* in the Late Eocene (35 Ma) with later entrance into Asia and North America in the Oligocene. This result was however, at odds with the fossil record reviewed in Chapter 3, which places the oldest fossil-remain of *Hypericum*, *H. antiquum*, in Siberia within the Eastern Palearctic region already in the Eocene.

The biodiversity we observe today is only a remnant of the one that once existed, since extinction might have erased some of the early divergent lineages. This implies that as we move to the distant past, inferences become more uncertain and less precise. It might happen that the geographic areas where a lineage now lives are different to those where its ancestors lived. For example, Rufhel (2011) found that the area of occurrence of the fossil *Paleoclusia* (North America), used for calibrating the stem-node of Clusiaceae, was outside the ancestral geographic range inferred for that node. A similar case is observed in *Hypericum* (above). Likewise, ancestral lineages do not necessarily share the same ecological requirements than their extant descendants, either because of dramatic changes in climate (e.g., the Holarctic region) or because evolution of new traits allows descendants to invade a completely new habitat (e.g., marine to freshwater). Phylogenetic or historical inference necessarily implies the assumption of some form of "actualism or uniformitarianism", i.e, the natural processes acting in the present are the same that those operating in the past (Hutton 1794; Gould, 1965; Lyell, 1830). Nevertheless, the fossil record might help us to escape this actualist "trap" by providing information on the potential geographic distribution and climatic tolerances of a lineage's ancestors.

In **Chapter 4**, we propose a new approach to incorporate fossil information, both ancestral ranges and climatic preferences, into the reconstruction of the spatio-temporal evolution of an organism. Fossils have traditionally been used in phylogenetic and biogeographic analyses to provide calibration points in estimating lineage divergence times (Ho & Philips, 2009). More recently, there have been attempts to explicitly include fossil geographic ranges into biogeographic analysis. These approaches require that the fossil lineage is included in the phylogeny, which often implies the coding of morphological characters (Mao et al. 2012) or rather arbitrary assumptions on the longevity of the fossil (Nauheimer et al., 2012). Here, we used instead the range of the fossil to constrain the inference of ancestral areas for the node to which the fossil is assigned. Besides the temporal and spatial aspects, fossils can also be used to reconstruct ancestral climatic tolerances (Nogués-Bravo et al., 2008; Maguire & Stigall

2009), but this approach has been so far limited to recent geological periods or small geographic regions. In Chapter 4, we extend this approach to the early history of the Cenozoic and used fossil-based ecological niche models to reconstruct the climate niche of *Hypericum* in the past, which when projected onto paleoclimate layers allows us detecting regions that were within the climate tolerance of the genus. These inferences were later used to inform biogeographic analysis, permitting us to detect regions that were outside the current geographic range of the group or that acted as potential "ecological dispersal corridors", allowing *Hypericum* lineages to move into new geographic regions (Donoghue, 2008). Finally, comparison between fossil-based and present climate optima informed us on the relative role of niche conservatism vs. niche evolution in the evolution of *Hypericum*.

Finally, in **Chapter 5,** we ask the central question: Which were the engines that fuelled the evolution of *Hypericum*? In particular, which were the factors or mechanisms explaining its extraordinary species richness compared to its closest relatives, i.e., other clusioid genera? This question bears on the interplay between niche dynamics and patterns of geographic evolution analysed in Chapters 1 and 4, and relates them with patterns of temporal diversification in the genus. Specifically, we tested two different hypotheses: the "key-innovation diversification" (KID) hypothesis states that a change in climatic tolerances to deal with the new temperature regimes in the Holarctic or the appearance of the herbaceous habit form have been "adaptive breakthroughs", allowing the genus to diversify rapidly and ultimately being responsible for its evolutionary success. In contrast, the "time-to-speciation effect" (TSE) hypothesis suggests that rather than "rapid bursts of diversification", the old age of the genus and steady accumulation of species through time can explain its large species richness. In this chapter, we used a wide array of methods, including ancestral character state reconstruction ENM geographic projections, and time-variable and trait-dependent macroevolutionary models to rule out between these two different evolutionary hypotheses. Palaeontology, Ecology and Evolutionary biology are disciplines that have traditionally been kept apart. One of the aims of this chapter, and throughout this doctoral work, is to show that the integration of all sources of evidence, being evolutionary relationships, the fossil record, or biogeographical and ecological reconstructions, can help us recover more realistic models of the past.

OBJECTIVES AND HYPOTHESES

The main objective of this thesis is to investigate patterns of biotic assemblage, with special focus on the effect of climate change and geological events on species diversity and distribution. We also aim to show the limitation of current approaches based on present diversity for this purpose and propose a method to integrate extinct (fossil) and ecological information in evolutionary analyses. To address these issues, we have studied the evolutionary history of the plant genus *Hypericum*.

Objectives Chapter 1: Bayesian inference of phylogeny, morphology and range evolution reveals a complex evolutionary history in St John's wort (*Hypericum*). This chapter has two general objectives: investigate evolutionary relationships in the genus *Hypericum* based on classic nuclear and chloroplast molecular markers and infer the evolutionary history (morphological and range evolution) of the group using statistical models. The underlying hypothesis is that present patterns of species diversity and distribution keep the signature of historical events. For this purpose, several particular objectives need to be accomplished:

- Infer phylogenetic relationships based on different molecular markers.
- Assess congruence between unlinked nuclear and plastid evidences.
- Assess congruence between plastid markers.
- Investigate the effect of missing data and partitioning strategies in molecular analyses.
- Evaluate and correct the estimation of branch lengths in Bayesian inference.
- Investigate the monophyletic status of *Hypericum* within the tribe Hypericeae.
- Compare the traditional classification of *Hypericum* with the new molecular results and assess the monophyly of the sections.
- Infer the evolutionary trajectories of some diagnostic morphological characters.
- Provide a temporal framework of cladogenetic events.
- Infer the dispersal history of the group, the area of origin, and main dispersal routes to colonize new continents, all within a frame of global change.
- Evaluate the potential of hierarchical Bayesian approaches in evolutionary inference.

Objectives Chapter 2: Utility of low-copy nuclear markers in phylogenetic reconstruction of *Hypericum* L. (Hypericaceae). The main objective of this chapter is to explore the potential of low copy nuclear markers (LCGs) to solve phylogenetic relationships in *Hypericum* in comparison with the commonly used nuclear ribosomal ITS region. The hypothesis is that LCGs have the potential to compensate the lack of resolution and low support values of classic nuclear and plastid markers. The particular objectives are:

- Screen several LCGs in a pilot study and select the best regions for phylogenetic reconstruction. The selection of new phylogenetic markers will be based on the appropriate levels of variation, lack of internal recombination and internal consistency of the phylogenetic signal.
- Design specific primers for amplification of these regions across different taxonomic groups in *Hypericum.*
- Asses variation, copy number and consistency for the selected LCGs, PHYC and EMB2765, in comparison with ITS.

Objectives Chapter 3: Paleobiology of the genus *Hypericum* (Hypericaceae): a survey of the fossil record and its palaeogeographic implications. The main objective of this chapter is to provide an overview of the fossil record of *Hypericum* to approximate the age of the genus and its geographic distribution in the past. The hypothesis underlying this chapter is that the fossil record contains not only temporal information, but also geographic, and can inform us on the past distribution of a lineage. The particular objectives are:

- Discuss the main diagnostic characters of seeds and pollen in *Hypericum*.
- Provide a survey of the fossil record of *Hypericum* and identify its oldest remain.
- Reassess the diagnostic characters of the oldest fossil record, *H. antiquum.*
- Synthetize the paleogeographic information provided by fossil sites locations.

Objectives Chapter 4: Using fossils to reveal the impact of ancient climate change in plant evolution. The objective of this chapter is to assess niche evolution and past distribution of *Hypericum* species. I also aim to present a new methodological approach based on the integration of biogeographic analyses and niche modelling techniques together with extant and extinct (fossil) data to infer the biogeographic history of the organisms. The underlying hypothesis is that extinction limits our ability to retrieve accurate biogeographic

patterns. Evolutionary analysis based only on present taxa might be biased towards the small fraction of diversity that survived until the present. Nonetheless, the fossil record provides a direct evidence of the extinct diversity, and therefore, its integration in the analysis might increase model realism and change our perception of the past. The particular objectives fulfilled in this chapter are:

- Compare the performance of Maximum Likelihood methods in biogeographic inference, which do not incorporate model uncertainty but instead allow the inference of widespread ancestral ranges, with the Bayesian approach presented in Chapter 1.
- Perform fossil-based ecological niche models (ENM) to infer potential areas of distribution in the past, ecological corridors specific for *Hypericum* during the Tertiary, and past climatic tolerances.
- Incorporate the temporal and geographic information provided by fossils into the biogeographic reconstruction.
- Incorporate a paleostratigraphic model into the biogeographic analysis that specifies the probability of connections between areas through time based on the configuration of continents.
- Incorporate a paleostratigraphic model into the biogeographic analysis that specifies the ecological connectivity (fossil-based ENM results) of *Hypericum* through time.
- Compare the accuracy of the biogeographic models constrained by past geological connections, fossil information and past ecological preferences with an unconstrained model.
- Infer the relative contribution of niche evolution vs. niche conservatism in the history of *Hypericum*.

Objectives Chapter 5: The strategy of the marathon runner: ancestral resilience drove the evolution of the species-rich, age-old genus *Hypericum* (St John's wort, Hypericaceae). In this chapter we aim to understand the mechanisms generating and maintaining diversity in *Hypericum*, and specifically to test the causal relationship between key-innovations, ecological or morphological, and cladogenesis. The underlying hypothesis is that key innovations promoted the origin and rapid diversification of *Hypericum*. To address this objective, we set up the following tasks:

- Infer the climatic niche preferences of the Hypericaceae stem ancestors of *Hypericum*.
- Infer the niche preferences of the ancestors of major lineages within *Hypericum.*

- Study diversification dynamics in the clusioid clade and within *Hypericum*; infer whether diversification rates have been constant or variable through time and across clades.
- Assess if putative rate shifts are correlated with the appearance of evolutionary novelties or specific abiotic factors.
- Assess if diversification rates are trait dependent, meaning that clade-specific characteristics, such as its geographic distribution or its habit form, confer some clades higher or lower speciation and extinction rates.
- Integrate these evidences to test the hypothesis that the evolution of new climatic tolerances and/or the appearance of new growth forms promoted the rapid diversification of *Hypericum*, which might explain its large species richness compared with its closest relatives.

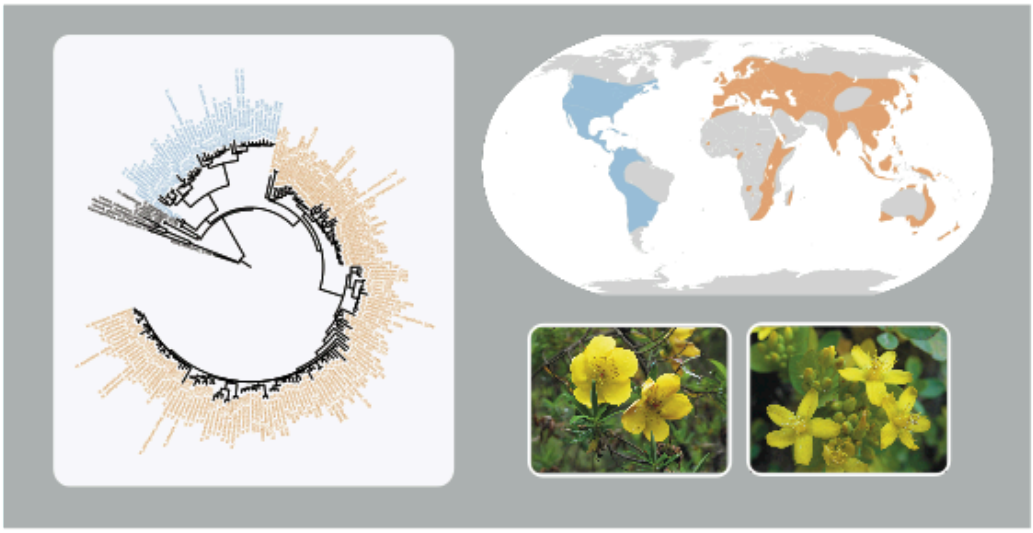
CHAPTERS

CHAPTER 1

Bayesian inference of phylogeny, morphology and range evolution reveals a complex evolutionary history in St. John's wort (*Hypericum*)

This chapter has been done in collaboration with Juan Jose Aldasoro and Isabel Sanmartín from the Department of Biodiversity and Conservation, Real Jardín Bótanico-CSIC, Spain

This chapter has been published in the peer-review journal *Molecular Phylogenetics and Evolution*, volume 67, pages 379 to 403 (2013)



CHAPTER 1

Bayesian inference of phylogeny, morphology and range evolution reveals a complex evolutionary history in St. John's wort (*Hypericum*)

Andrea Sánchez Meseguer, Juan Jose Aldasoro, Isabel Sanmartín

A b s t r a c t

The genus *Hypericum* L. (''St. John's wort'', Hypericaceae) comprises nearly 500 species of shrubs, trees and herbs distributed mainly in temperate regions of the Northern Hemisphere, but also in high-altitude tropical and subtropical areas. Until now, molecular phylogenetic hypotheses on infrageneric relationships have been based solely on the nuclear marker ITS. Here, we used a full Bayesian approach to simultaneously reconstruct phylogenetic relationships, divergence times, and patterns of morphological and range evolution in *Hypericum*, using nuclear (ITS) and plastid DNA sequences (*psbA-trnH, trnS-trnG, trnL-trnF*) of 186 species representing 33 of the 36 described morphological sections. Consistent with other studies, we found that corrections of the branch length prior helped recover more realistic branch lengths in by-gene partitioned Bayesian analyses, but the effect was also seen within single genes if the overall mutation rate differed considerably among sites or regions. Our study confirms that *Hypericum* is not monophyletic with the genus *Triadenum* embedded within, and rejects the traditional infrageneric classification, with many sections being paraor polyphyletic. The small Western Palearctic sections *Elodes* and *Adenotrias* are the sister-group of a geographic dichotomy between a mainly New World clade and a large Old World clade. Bayesian reconstruction of morphological character states and range evolution show a complex pattern of morphological plasticity and inter-continental movement within the genus. The ancestors of *Hypericum* were probably tropical shrubs that migrated from Africa to the Palearctic in the Early Tertiary, concurrent with the expansion of tropical climates in northern latitudes. Global climate cooling from the Mid Tertiary onwards might have promoted adaptation to temperate conditions in some lineages, such as the development of the herbaceous habit or unspecialized corollas.

1. Introduction

Bayesian inference techniques have become very popular in phylogenetics because of the relative ease with which these techniques allow biologists to infer evolutionary patterns using complex and realistic models (Ronquist, 2004). Markov Chain Monte Carlo Bayesian approaches have now been developed to answer evolutionary questions, ranging from the time and place of origin of lineages to inferring the evolution of morphological traits, while accounting for phylogenetic and model uncertainty (Drummond and Rambaut, 2007; Huelsenbeck and Bollback, 2001; Lemey et al., 2009; Ronquist and Sanmartín, 2011; Sanmartin et al., 2008). Here, we use this full Bayesian approach (Ronquist, 2004) to simultaneously reconstruct phylogenetic relationships, lineage divergence times and ancestral areas in the old worldwide-distributed plant genus *Hypericum* (Nürk and Blattner, 2010; Robson, 1981), while integrating out uncertainty concerning tree topology and other model parameters. *Hypericum* L. represents one of the 100 largest angiosperm genera of the world (Carine and Christenhusz, 2010), with over 496 species (including other Hypericaceae genera (Nürk et al., 2012), or 500 in the most recent Robson's (2012) revision) of trees, shrubs and herbs. The genus is distributed in almost every continent and ecosystem, being absent only in the poles, arid deserts, and low-altitude tropical areas (Fig. 1) (Robson, 1977). *Hypericum* is a relatively old genus as suggested by its fossil record dating back to the Early–Mid Tertiary, ca. 37–34 Ma (Meseguer and Sanmartín, 2012). Some Hypericum species, such as *Hypericum perforatum* L. (common St. John's wort), are economically important in pharmacology because of their active compounds hypericine and pseudo-hypericine, which are used as painkillers, antidepressants or anticancer treatments (Barnes et al., 2001). In this aspect, a phylogenetic hypothesis for the genus Hypericum could be interesting for bioprospecting.

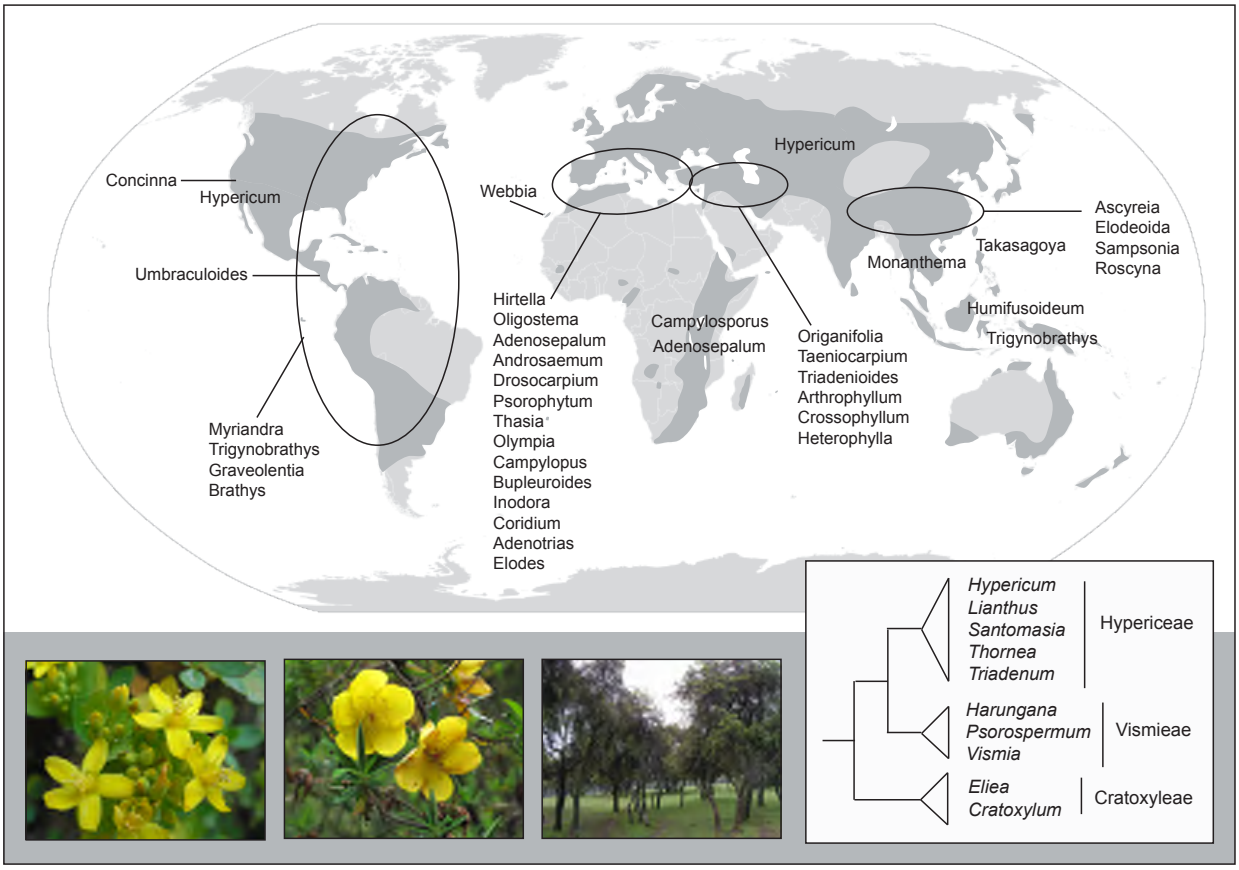


Fig. 1. Present distribution of *Hypericum* species. Map showing the current distribution of *Hypericum* species (modified from Robson, 1977); for each section the regions harboring the highest number of species are given. Inset: Schematic representation of phylogenetic relationships among the genera of family Hypericaceae, showing division into tribes. Below, from left to right pictures of *H. tortuosum* flowers (section *Triadenioides*), leaves and flowers of *H. revolutum* (Campylosporus) and habit of *H. revolutum*.

Current Angiosperm classification (APGIII 2009, Stevens, 2007) includes the genus *Hypericum* in the family Hypericaceae, belonging to the large clade of mostly tropical plants known as the ‘‘clusioid clade’’ (Davis et al., 2005; Gustafsson and Persson, 2002; Ruhfel et al., 2011; Wurdack and Davis, 2009). Three tribes are recognized within Hypericaceae: the tropical tribes Vismieae Choisy (*Vismia* Vand., *Harungana* Lamarck and *Psorospermum* Spach) and Cratoxyleae Bentham & J.D. Hooker (*Cratoxylum* Blume, *Eliea* Cambess.), and the widespread tribe

Hypericeae Choisy, including the genera *Triadenum* Raf., *Thornea* Breedlove & McClintock, *Santomasia* N. Robson, *Lianthus* N. Robson, and *Hypericum* (Fig. 1, inset). Yet, relationships among genera remain unclear (see below).

Hypericum is one of few large genera with an almost complete taxonomic treatment. Robson (Robson, 1977, 1981, 1985, 1987, 1990, 1996, 2001, 2002, 2006, 2010a, 2010b, 2012) published a series of monographs in which he described numerous species and defined the main diagnostic characters for the taxonomy of the genus. Robson divided the genus into 36 sections (see Nürk and Blattner (2010) for a synthesis of Robson’s classification), and proposed relationships between sections based on the evolutionary direction of certain traits, such as the habit form, presence of dark glands, corolla shape, or the number of stamen fascicles. Based on Robson’s study, Nürk and Blattner (2010) carried out the first morphological cladistic analysis of this genus, and concluded that some of these diagnostic characters were under convergent evolution. They also found discrepancies with Robson’s sectional classification, and suggested the inclusion of the monotypic genus *Santomasia* within *Hypericum*.

In contrast to morphological studies, work at the molecular level has been slower in *Hypericum*, probably due to the difficulty to work with such a large and cosmopolitan genus. Ruhfel et al. (2011) analyzed relationships beyond the genus level in the clusioid clade and concluded that *Hypericum* is not monophyletic, with genera *Santomasia*, *Triadenum*, and *Thornea* embedded within. However, their study included only 21 *Hypericum* species, so little could be inferred in terms of infrageneric relationships. Other molecular studies focusing on

interspecific relationships in *Hypericum* were too limited in both taxonomic and geographic coverage (Crockett et al., 2004; Heenan, 2008; Park and Kim, 2004; Pilepić et al., 2011). Just recently, Nürk et al. (2012) published the first deep-sampled molecular phylogeny for the genus including ca.40% of the species diversity. They confirmed the inclusion of *Triadenum* within *Hypericum*, but, contrary to Ruhfel et al. (2011), recovered *Thornea* as the sister group of *Hypericum*. They also reconstructed ancestral states for some diagnostic characters, confirming many of Nürk and Blattner's (2010) conclusions. All the above-mentioned species-level phylogenies were based solely on ribosomal nuclear internal transcribed spacer. It is well known, that phylogenies based on ITS alone can be problematic because this marker displays a complex evolutionary behavior owing to concerted evolution among its multiple copies (Álvarez and Wendel, 2003). Also, biological processes such as hybridization, duplication, introgression, or incomplete lineage sorting may obscure the correlation between gene trees and the species tree. Thus, additional inclusion of plastid genes is desirable when reconstructing evolutionary relationships among species (Doyle, 1992).

Hypericum is unique within the clusioid clade in its variable habit form and mainly temperate distribution (most of the other genera are woody elements of tropical forests). The largest diversity of the genus is found in temperate areas of the Northern Hemisphere, Eurasia and North America, but some sections have reached high-altitude areas in the tropical regions, such as the South American Andes, where they exhibit some remarkable radiations, i.e., the 88 species in section *Brathys* (Robson, 2012). In Africa, the genus exhibits an interesting biogeographic disjunction, in which related species are distributed along the margins of the continent (e.g., Macaronesia, the Eastern African Mountains, and South Africa) as well as in Madagascar, in what has been called the ''Rand Flora pattern'' (Sanmartín et al., 2010) (see Fig. 1). Interestingly, Robson (1981) placed the origin of *Hypericum* in Africa and hypothesized that the character states exhibited by the Afromontane species, such as the treelet habit or presence of dark glands, were the ''ancestral'' states for the genus. He thought that the genus was very old and probably originated before direct land connections between Africa and the other parts of Gondwanaland broke off in the Mesozoic. This hypothesis, however, is at odds with a recent revision of the fossil record of *Hypericum* (Meseguer and Sanmartín, 2012), and with molecular estimates of divergence times in Malpighiales (Davis et al., 2005), dating the split between Hypericaceae and its sister family Podostemaceae in the Early Paleocene. Meseguer and Sanmartín (2012) placed the oldest fossil evidence of *Hypericum* in the Late Eocene of West Siberia, and suggested that the ancestors of the genus were part of the boreotropical forest belt that covered the Holarctic during the warm periods of the Early Tertiary (Tiffney, 1985a; Wolfe, 1975). Other studies (Nürk and Blattner, 2010; Nürk et al., 2012) have placed the origin of the genus in the Mediterranean Region based on the basal position of the Mediterranean clades. However, none of these hypotheses were tested within a formal biogeographic analysis.

In this study, we present the first species-level phylogeny of *Hypericum* based on both biparentally inherited nuclear DNA (nrDNA) and maternally inherited plastid DNA (cpDNA), and covering 40% of the described species and 33 out of the 36 proposed morphological sections. We use the full hierarchical Bayesian approach described in Huelsenbeck and Bollback (2001) and Ronquist (2004) to reconstruct the evolution of some of the most variable and taxonomically important characters in the genus. Finally, we applied a Bayesian discrete phylogeographic model (Lemey et al., 2009) in conjunction with relaxed clock dating and fossil evidence to estimate ancestral areas and the main migration events within the history of the genus.

2. Materials and methods

2.1. Taxon sampling

Sampling effort was aimed to cover morphological and geographic variation of the genus. We sampled ca. 40% of the species (186 species out of 496) and more than 90% of the sectional variation (33 sections out of 36). Our sampling is comparable with that of Nürk et al. (2012), which included 200 species, nearly 70% of them represented in this study. However, our final dataset comprises 3032 characters, a fourfold increase over similar studies at infrageneric level on *Hypericum* (Crockett et al., 2004; Park and Kim, 2004; Nürk et al., 2012), all of which were based on nuclear ITS. Missing sections were the East Mediterranean section *Origanifolia* with 13 species, and the monotypic sections *Concinna* (N. America) and *Umbraculoides* (Mexico). The missing species mostly belong to the large sections *Brathys* and *Trigynobrathys* from America, *Ascyreia* from Asia and *Hirtella* and *Taenioarpium* from Levant. We made a special effort to increase the sampling of African sections, which were usually poorly represented in previous phylogenetic studies of *Hypericum*. We also

included representatives of other Hypericaceae genera: *Triadenum* and *Thornea*, the latter only represented by GenBank ITS accessions, from tribe Hypericeae; *Vismia* and *Harungana* representing sister-tribe Vismieae, and genus *Eliea* from tribe Cratoxyleae. The latter was used as the most external outgroup to root the trees, following previous studies (Ruhfel et al., 2011; Wurdack and Davis, 2009). DNA data was obtained from fresh material collected in the field and preserved in silica gel, and from dry material preserved at several herbaria (Appendix A). GenBank accessions from previous studies of *Hypericum*, mostly ITS, were also included in the final dataset. Species names, voucher information and Genbank (NCBI dataset) accession numbers are shown in Appendix A.

2.2. DNA extraction, amplification and sequencing

Three plastid (*trnS-trnG, psbA-trnH, trnL-trnF*) and one nuclear (ITS) region were amplified using universal and newly designed primers. The intergenic spacers (IGS) *trnS-trnG* and *psbA-trnH* were amplified using primers from Hamilton (Hamilton, 1999). Additional degenerate internal primers were designed for *trnS-trnG: trnSG-A* (5' - ACT GCT TCG ACT MAA TTT MG-3') and *trnSG-B* (5' AGG ATT MGG ATT GMT CTT GTT TC-3') using the software OligoCalc (Kibbe, 2007). We amplified the *trnL-trnF* region using primers c–f from Taberlet et al. (1991). The ITS region was amplified using the universal primers ITS4 and ITS1a (Aguilar et al., 1999; White et al., 1990). For some species that were difficult to amplify, we used also the internal primers ITS2 and ITS3 (White et al., 1990). DNA was extracted from leaf tissue samples using the QIAGEN DNeasy plant kit (Qiagen, Hilden, Germany) at the laboratories of the Real Jardín Botánico-CSIC (Madrid, Spain), and following the manufacturer's protocol. Amplification was achieved in a 25 µl reaction volume using the PCR mix BioMix (Bioline, Germany). The PCR cycling conditions were as follows: 95 ºC for 5 min, 35 cycles of [94 ºC for 30s, 52–56 ºC for 1 min, 72 ºC for 1.5 min] and a final extension step of 10 min at 72 ºC. PCR products were checked on 1% agarose gels and sequencing was performed at Macrogen, Inc. (Seoul, South Korea), using the initial PCR primers. Amplified products were purified using the Qiagen PCR Purification Kit. We occasionally got multiple fragments of different lengths, especially in *psbA-trnH*, which were directly isolated from the gel using the Zymoclean Gel DNA Recovery kit (California, USA).

In most cases we obtained unambiguous sequences, but some ITS sequences showed more than one polymorphic site (e.g., clear double peaks in both sequence strands). To screen for possible variants, PCR products were cloned using the CopyControl cDNA, Gene and PCR Cloning Kit (Epicentre, Madison, USA), according with the manufacturer's manual. Fifteen positive colonies were selected and amplified using the universal primers T7 and pCC1/pEpiFOS RP-2 reverse sequencing primer. No sequences with >5% divergences were found among the clones, so we included these sequences in the final dataset.

2.3. Phylogenetic methods

DNA sequences were edited using Sequencher 4.7 (Gene Codes, Ann Arbor, MI). High levels of sequence variation, especially in relation to the presence of indels or gaps, were found in all markers, in agreement with other studies of Malpighiales (Davis et al., 2005; Wurdack and Davis, 2009). Thus, sequence alignment was difficult and we followed a three-stage approach. First, sequences were aligned using the online version of MAFFT v.6 (Katoh and Toh, 2008), with the default option L-INS-I (Katoh et al., 2002; Katoh and Toh, 2008), and visually adjusted using the software Se-Al v2.0a11 Carbon (Rambaut, 2002). Second, the software Gblocks v.0.91b (Castresana, 2000) was used to identify and remove ambiguously aligned regions such as large segments of non-conserved positions or with a large density of gaps. We used this approach only for the ITS marker, because alternative analyses with owithout Gblocks showed that including these ambiguous regions in the chloroplast alignments yielded stronger statistical support for several clades. Third, ''informative'' gaps were coded as binary characters using the ''simple gap'' coding (Simmons and Ochoterena, 2000) implemented in the software SeqState version 1.4.1 (Müller, 2005). Although gaps are a potential source of information in phylogenetic analysis, they can be difficult to align and might artificially increase the homoplasy in the dataset. We only coded gaps as informative characters if they could be unambiguously aligned across species, such as positionally homologous deletions embedded within an otherwise conserved segment. This was the case of the *trnS-trnG* and *trnL-trnF* markers, where gaps grouped clades that were also supported by standard substitution characters. Conversely, gaps were coded as missing data (non-informative) in the *psbA-trnH* dataset — or removed with Gblocks prior to analysis in ITS— because they could not be unambiguously aligned and including them lowered general clade support values.

2.4. Phylogenetic analysis

2.4.1. Single-marker and combined analyses

We used Bayesian inference (BI) implemented in MrBayes v3.2 (Ronquist et al., 2012) to infer phylogenetic relationships in Hypericum. Substitution models for each gene were selected based on the Akaike Information Criterion (Akaike, 1973) implemented in MrModeltest 2.3 (Nylander, 2004). The GTR model with rate variation among sites following a gamma distribution (GTR+G) was the best model for the chloroplast markers, and the same model but with a proportion of invariable sites (GTR+G+I) was selected for the ITS marker. For the gap partition in *trnS-trnG* and *trnL-trnF*, we applied a restriction site model (F81) with ''lset coding = variable'' to accommodate the ascertainment bias. Two independent runs of three Metropolis-coupled chains each were run for 10–20 million generations, sampling every 1000 generations. Mixing and convergence among chains were assessed using the standard deviation of split frequencies in MrBayes and the effective sampling size criterion (values >200) in Tracer v1.6 (Rambaut and Drummond, 2009). We also used the online tool AWTY (Nylander et al., 2008) to monitor cumulative posterior probabilities and among-run variability of split frequencies to ensure that all chains have reached the same stationary phase. After discarding the first 1–2 million generations (10–20% of samples) as ''burn-in'', the remaining samples from the independent runs (approx. 18,000–16,000) were summarized into a 50% majority rule consensus tree with clade posterior probabilities to approximate the posterior distribution of the phylogeny. To speed up convergence, we estimated a maximum likelihood tree with the fast software RAxML v.7.2.8 online version (Stamatakis et al., 2008), and employed this tree as the starting value (''starting tree'') for the tree parameter (tau) and the branch length parameter (V) with the MrBayes v3.2 commands: ''startvals tau = mystarttree V = mystarttree''. To avoid using the same starting tree in the two independent runs, which makes it more difficult to detect convergence problems, we introduced random perturbations in the ML tree with the command ''mcmcp nperts = 0.1''; we then used these slightly perturbed versions of the original tree as starting trees for the two runs. Additionally, we used the program GARLI v2.0 (Zwickl, 2006), which performs highly efficient likelihood searches, to estimate the phylogeny under the maximum likelihood criterion, using the evolutionary model selected by MrModelTest, and repeating the analysis twice starting from different random trees. Clade support was assessed by non-parametric bootstrapping using 500 replicates in GARLI.

Before concatenating the different genes into a single dataset, we assessed congruence by running analyses on each individual marker, and comparing the resulting consensus trees for cases of ''well-supported conflicting clades'', i.e., clades that are significantly supported (>95 Bayesian posterior probability) in one gene tree but not in the consensus trees of the other markers. We also tested for substitutional saturation in each marker by plotting the uncorrected pairwise sequence distances (''p'') against ML distances derived in PAUP v4.0b10 (Swofford, 2002) under the selected nucleotide model, and checking for deviation from linearity of plots. Since no significant incongruence was found among the plastid markers (but see below), we combined them into a single dataset using the program Phyutility v2.2 (Smith and Dunn, 2008), which was analyzed in MrBayes under the same settings as above. The ITS marker was analyzed separately to compare topologies between the nuclear and plastid genomes and to avoid artifacts derived from combining markers with different levels of heterogeneity in mutation rates.

2.4.2. Missing data and partitioning strategy

Sensitivity analyses were carried out to evaluate the effect of missing data and different partitioning strategies in the combined three-marker cpDNA dataset. Missing data, due to failure to amplify some markers for certain specimens, may introduce problems in Bayesian phylogenetic inference (Lemmon et al., 2009; Simmons, 2012) but see (Wiens, 2006; Wiens and Morrill, 2011) for a different view). To evaluate the effect of the missing data in our cpDNA dataset, we run Bayesian and ML analyses using the same parameter settings as above on three different concatenate matrices: (a) ''No-missing'': including only those specimens that were represented in all three chloroplast markers; (b) ''Two-markers'': including only those specimens sequenced for at least two markers; (c) ''All-specimens'': including all sequenced specimens (approximately 53% of specimens missing at least one marker). We then compared the resulting trees from these analyses in terms of tree topology, clade support, and level of resolution, i.e., number and percentage of resolved nodes over the total number of nodes for a tree of this size. Results showed that the presence of missing data decreased the level of resolution in the resulting phylogeny: ''Nomissing'': 79 resolved clades (87% over total number); ''Two-markers'' 112 (77%); ''All-specimens'': 119 (62%). The overall topology and major clades were recovered by all three datasets. Because the ''All-specimens'' dataset contains more data, phylogenetic discussion will be based on this. However, clade

support and resolution are lower than in the ''Two-markers'' dataset, so we used the latter for the reconstruction of ancestral states and the biogeographic-dating analysis.

We also performed a sensitivity analysis to evaluate the impact of different partitioning strategies. The benefits of creating partitions – assigning an independent evolutionary model to each molecular marker in a multi-gene Bayesian analysis – have been discussed in several studies (Marshall, 2010; Marshall et al., 2006; Nylander et al., 2004). Partitioning, especially if allowing the overall mutation rate to differ among markers, can improve the fit to the data and decrease the variance, which results in higher clade support values and more accurate phylogenetic relationships (Marshall et al., 2006; Nylander et al., 2004). Yet, recent studies (Brown et al., 2010; Marshall, 2010) have warned about the dangers of a partitioned multi-gene dataset when the rate of mutation differs highly among partitions. When data from independent partitions evolve at very different rates, the analysis can get trapped in regions of low posterior density and ''overly'' long trees, where branch lengths are severely overestimated. One solution to this problem is to increase the value of the λ parameter that controls the exponential prior on branch lengths (1/λ), which has the effect of pushing up the exponential prior more tightly around small branches (Brown et al., 2010; Marshall, 2010; Marshall et al., 2006). To test this effect in our concatenate cpDNA ''All-specimens'' dataset, we run Bayesian analyses with three different partitioning strategies: (1) ''All-unpartitioned'' dataset, in which a single substitution model was applied to all sites; (2) ''All-partitioned uncorrected'' dataset in which "rate multipliers" m_1, m_2 . . m_n were estimated per partition to accommodate rate variation (''prset ratepr = variable") but the branch length prior was assigned the default value (λ= 10, 1/λ = 0.1); and 3) ''All-partitioned corrected'' dataset accommodating among-partition rate variation (''prset ratepr = variable''), but lowering the value of the exponential prior (λ = 100, 1/λ = 0.01) using the command ''prset brlenspr = Unconstrained:Exp(100)''. Bayes Factors, based on the harmonic mean of the two runs (Kass and Raftery, 1995), were used to compare the marginal likelihood and fit to the data of each partitioning strategy.

2.5. Ancestral state reconstruction

Bayesian ancestral state reconstructions (ASRs) were performed in MrBayes v 3.2 on the concatenate ''Two-markers'' chloroplast dataset using the full hierarchical Bayesian approach, i.e., integrating out uncertainty concerning tree topology and other model parameters (Huelsenbeck and Bollback, 2001; Ronquist, 2004). We did not use the ITS dataset because higher rate heterogeneity and recombination in nuclear markers may hinder the estimation of evolutionary rates and associated branch lengths (Álvarez and Wendel, 2003). This makes ITS less appropriate for inferring ancestral states and lineage divergence times, especially if as in Hypericum there are changes in life history traits: e.g., shifts between woody/perennial and herbaceous habits (Kay et al., 2006; Litsios and Salamin, 2012). We reconstructed evolutionary patterns in seven morphological diagnostic traits: habit form, presence of dark glands, number of fasciclodes (vestigial fascicles), ornamentation of seed testa, shape of flower corolla, and number and degree of fusion of stamen fascicles; see Supplementary information (SI) Appendix for a description of characters. Some species were coded as polymorphic for certain characters, e.g., H. revolutum exhibits both the cyathiform and stellate corollas (Fig. 4), which is interpreted as ambiguity in Bayesian ASR. We reconstructed ancestral states in eight lineages representing the main clades recovered in the phylogenetic analyses, which also received high clade support (>95% except for clade C). ach morphological character was added to the end of the molecular matrix and modeled according to the Mk1 model of Lewis (Lewis, 2001) (standard discrete model), with its own partition-specific rate multiplier. We analyzed each matrix (plastid dataset + 1 character) separately to minimize the influence of morphology in the estimation of phylogenetic relationships. All other settings were identical to those used above in the Bayesian inference of the phylogeny (e.g., by-gene partitioned analysis with corrected lambda prior, ML starting tree).

2.6. Molecular dating

Absolute lineage divergence times in *Hypericum* were estimated in BEAST (Drummond and Rambaut, 2007) using a Bayesian relaxed clock-model. The chloroplast ''Two-markers'' dataset was used for the analysis with the following settings: a by-gene partitioned dataset with GTR+G as substitution model, Yule tree prior, and uncorrelated lognormal relaxed clock (UCLD). Bayes Factors were used before to discriminate between different model clocks (strict/relaxed) and partitioning strategies (partitioned/unpartitioned). Topological constraints were enforced to include prior phylogenetic knowledge in the analysis. In particular, initial BEAST runs did not recover the

sister group relationship between *H. elodes* and *H. aegypticum* with the rest of *Hypericum* (supported by Nürk et al. (2012) and our MrBayes analyses), or the position of Eliea as sister to Vismieae–Hypericeae, which is also supported by Rufhel et al.'s (2011) clusioid clade phylogeny. These relationships were enforced in all subsequent BEAST analyses. To avoid conflict between the starting tree and the topological priors in the analysis, we used the ''allcompat'' tree from the Bayesian analysis, with branch lengths calibrated by Non-Parametric Rate Smoothing (NPRS) (Sanderson, 1997) using the software TreeEdit v.1.0a10 (Rambaut and Charleston, 2001) and a fixed age for the root node calibration (see below). Two replicate MCMC searches of 30 million generations each were run under these settings and their results pooled using the software LogCombiner v. 1.7.2 (after removing 25% samples as burn-in). We used Tracer 1.6 to determine stationarity of the Markov chain and to verify that all parameters have effective sampling sizes (ESSs) >200. TreeAnnotator v1.4.8 (Drummond and Rambaut, 2007) and FigTree v. 1.3.1 (Drummond and Rambaut, 2007), respectively, were used to generate and visualize the resulting maximum clade credibility (MCC) tree.

We used two external calibration points based on fossil evidence to obtain absolute divergence times:

(a) The root node, the crown age of Hypericaceae or the split between *Eliea* and the rest of the tree, was constrained according to Ruhfel (2011). He dated a molecular phylogeny of the clusioid clade (Ruhfel et al., 2011) using two fossil calibration points: the Upper Cretaceous macrofossil *Palaeoclusia chevalieri* and the Eocene pollen fossil *Pachydermites diederexii*. The fossil *Pachydermites* is placed with confidence as the most recent common ancestor (MRCA) of Pentadesma and Symphonia (Ruhfel, 2011). However, the phylogenetic position of *Palaeoclusia* is still controversial. Ruhfel (2011) conducted two independent analyses with different positions of the fossil in the phylogeny: as the stem age of the clusioid clade (OC position: MRCA of Ochnaceae s.l. and the clusioid clade), and as the stem node of the Clusiaceae family (BC position: MRCA of Bonnetiaceae and Clusiacae s.s.). Depending on the position of *Palaeoclusia*, he obtained a crown age for Hypericaceae between 58.9 and 71.5 Ma (OC and BC, respectively). To integrate this uncertainty in our analysis, we assigned a normal prior to the crown age of Hypericaceae, with mean 65.2 Ma (the mean of the BC and OC ages) and a std. of 11 to span the entire confidence interval (47.9–86.4 Ma) obtained by Ruhfel (2011)).

(b) To constrain the crown age of *Hypericum*, we used the fossil seed *Hypericum antiquum*, from the Late Eocene of West Siberia (Arbuzova, 2005), considered the oldest fossil remain of the genus (Meseguer and Sanmartín, 2012) (see SI Appendix for a discussion on the phylogenetic position of the fossil in our phylogeny). We used a lognormal prior to reflect the uncertainty in the fossil calibration (as recommended by Ho and Phillips, 2009), with the uppermost limit of the time interval (Priabonian) as a minimum hard bound (offset = 33.9 Ma) and a standard deviation (Std = 0.7) that includes the entire geological interval (33.9–37.2 Ma) (Walker and Geissman, 2009).

2.7. Biogeographic analysis

We inferred posterior estimates of ancestral ranges for the main lineages in the phylogeny in two different ways. First, we use Bayesian ASR and a similar approach to the morphological reconstruction above. Geographic distribution was coded as a multistate character and added to the ''Two-marker'' dataset as a standard morphological partition using the morphological discrete Mk1 model. Seven discrete areas were defined according to the paleogeographic history of the continents (see Fig. 5 and SI-Appendix): eastern Palearctic (EP), western Palearctic (WP), Nearctic (Ne), Neotropical (Nt), Afrotropical (AF), Oceania (OC), and IranoTuranian–Himalayan region (ITH). Ancestral ranges were estimated for the eight clades described above. Second, we used the Bayesian discrete phylogeographic approach of Lemey et al. (2009), implemented in BEAST v.1.6.2, to infer ancestral ranges and trace the history of geographic movement across regions in *Hypericum*. In the Bayesian ASR, branch lengths are measured as expected number of substitutions per site per unit of time, as in a phylogram. Although this is appropriate for inferring the rate of morphological evolution, especially if there are associated changes in life history traits (Litsios and Salamin, 2012), time-calibrated branch lengths measured as units of absolute time (as in a chronogram) are probably more interesting for inferring biogeographic history because dispersal barriers arose and fell through time (Ree and Sanmartín, 2009). Lemey et al.'s (2009) biogeographic method allows jointly estimating the posterior distribution of topologies, divergence times, and

ancestral ranges given molecular data and the geographic location of each species. The model is very similar to the Bayesian Island Biogeography (BIB) model described in Sanmartin et al. (2008) in that movement between geographic areas is modeled as a discrete-state continuous-time Markov chain (CTMC) with transition states (ancestral ranges) limited to single areas (Ronquist and Sanmartín, 2011). Dispersal rates between areas and ancestral ranges at nodes are estimated using MCMC Bayesian inference (Lemey et al., 2009). We run two replicate searches of 30 million generations, using uninformative priors for dispersal rates instead of constraining them by geographic distance (Lemey et al., 2009), since this changed over time with continental movement; the remaining BEAST settings were identical to the ones described in ''Molecular dating''. The discrete CTMC model implemented in BEAST v.1.6 can only handle single-area terminals. Because we used such all-encompassing areas (i.e., continents or major continental landmasses), most terminals ended up being endemic to a single operational area (Nearctic, Africa, etc.). As a result, there were only seven widespread species in our dataset, i.e., occurring in more than one region (SI-Appendix). We coded those widespread terminals as occurring in the area where the voucher was collected. However, this could introduce bias in the analysis if the sampling was not homogeneous among regions or the terminals represent larger clades with a widespread distribution such as outgroups. To examine the influence of forcing widespread terminals to occur in single areas, we carried out a second analysis in which these terminals were coded for the alternative area, for example, *Vismia* was coded as South American instead of African (see Fig. 5 and SI-Appendix).

3. Results

3.1. DNA sequence variation

Table 1 summarizes the main characteristics of the genomic regions studied. In total, 669 sequences were analyzed, of which 587 were generated in this study. The ITS dataset yielded a matrix of 520 characters and 252 specimens. The combined matrix of chloroplast regions (''All-specimens'') has 3072 aligned positions and 192 taxa. The saturation plots for the individual markers show a strong fit to a linear regression, although ITS and *psbA-trnH* present the lowest correlation and their saturation plots indicate slight levels of substitutional saturation at the deeper divergences (see Table 1 and SI Fig. 1). All data matrices can be obtained on request from the corresponding author.

3.2. Topological congruence and sensitivity analysis

Figs. 2 and 3 show the Bayesian consensus trees with BI and bootstrap values obtained with GARLI for ITS and the combined ''All-specimens'' cpDNA dataset, respectively; consensus trees for each individual chloroplast marker, *psbA-trnH, trnS-trnG, trnL-trnF*, are shown in SI Fig. 2. Overall, there was general topological congruence among plastid markers, with the exception of some cases of well-supported incongruence affecting *psbA-trnH* (SI Fig. 2). One conflict concerns several species from sections *Hypericum*, *Adenosepalum* and *Crossophyllum* that form a clade in *psbA-trnH*, but are scattered along the tree in the other cpDNA markers (SI Fig. 2 and Appendix A). Another relates to the placement of several not closely related specimens (e.g., *H. balearicum_C40, H. coris_C23, T. petiolatum_C16, H. synstylum_C11*) that occupy different positions in *psbA-trnH* than in all other markers (SI Fig. 2 and Appendix A). We discarded human error by repeating the sequencing of these specimens, and ensuring that they fall in the same position than in the first analysis. Many of these relationships are not supported by the traditional classification based on morphological characters (Robson, 1977) and do not appear in the ITS tree. Moreover, analyzing the combined plastid dataset with (SI Fig. 3) and without these incongruent sequences (Fig. 3) did not affect the overall topology of the tree, which recovered the same major groupings. Excluding *psbA-trnH* altogether – analyzing a combined matrix with *trnS-trnG* and *trnL-trnF* alone (SI Fig. 4) – also recovered a tree topology and groupings similar to Fig. 3, although including all three chloroplast markers increased significantly the support for many individual clades. Therefore, in discussing phylogenetic relationships in *Hypericum*, we used the complete (three markers) cpDNA dataset (Fig. 3) but excluding the problematic *psbA-trnH* sequences. Comparison between the combined cpDNA phylogeny (Fig. 3) and the ITS tree (Fig. 2) showed general levels of congruence, with all major clades supported by the two genomes. There was generally lower support in the ITS tree compared to the cpDNA tree, but there were a few cases of well-supported conflict (>95 pp) affecting species-level relationships. For example, the position of several species of the section *Adenosepalum* varies between the ITS and cpDNA trees; other species are assigned to different clades such as *H. scouleri* or *H. monanthemum* (Figs. 2 and 3).

Table 1
Sequence characteristics of the different nrDNA and cpDNA regions. Sequence variation and characteristics of the chloroplast regions *psbA-trnH, trnL-trnF and trnS-trnG*, and the nuclear intergenic spacer ITS with and without the ambiguously aligned regions (excluded with the software Gblocks: ''ITS Gblocks'').

	psbA-trnH	*trnL-trnF*	*trnS-trnG*	ITS (Gblocks)	ITS
Number of accessions	142	173	108	252	252
Aligned length	1322	727	1023	520	783
Un-aligned length[a]	525	604	670	518	710
Indel characters (%)	797 (60.3)	123 (17)	353 (30)	2 (0.38)	73 (9.3)
Constant characters	865	468	619	193	381
Parsimony-uninformative characters	125	89	134	72	78
Parsimony-informative characters (%)	332 (25.1)	170 (23.4)	270 (26.4)	255 (49)	324 (41.3)
Mean sequence divergence[b] (%)	0.34–0 (5.14)	0.37–0 (4.99)	0.28–0 (4.18)	0.78–0 (11.61)	0.75–0 (13.32)
Saturation (r^2 values)	0.987	0.997	0.99	0.986	0.98

[a] Total unaligned length per marker was obtained by averaging the length of 10 sequences per marker.
[b] Mean sequence divergence (%) estimated in PAUP over the total number of sequences.

Sensitivity analysis showed that the All-specimens ''All-partitioned'' datasets fit the data significantly better than the ''Allunpartitioned'' analysis (Table 2). Moreover, the ''partitioned uncorrected'' analysis with default branch length priors resulted in Bayesian consensus trees that were several orders of magnitude longer than the ML trees. In fact, the 95% credibility intervals of the Bayesian branch length estimates did not include the ML branch estimates, something that has been interpreted as evidence of inaccurate branch length estimates in MrBayes (Brown et al.,2010). By contrast, the ''All-partitioned corrected'' analysis with a lower exponential branch length prior resulted in very similar average branch length values between the BI and ML methods (Table 2). Interestingly, the same behavior was observed when we introduced the lambda correction in the single-gene analyses, resulting in average branch lengths that were shorter and more similar to the ML values (Table 2). The latter also resulted in a speed up in convergence among runs and better estimates for the among-site rate variation gamma parameter. One possible explanation is that considerable rate heterogeneity exist not only among partitions but also among sites within partitions, especially in ITS, where highly conserved regions are followed by long segments of variable, non-conserved positions. Therefore, all results presented here, are based on the corrected branch-length analyses (''All-partitioned corrected'' strategy).

3.3. Phylogenetic relationships

The combined cpDNA and ITS phylogenies show Vismieae as sister group to *Hypericum*, which is recovered as non-monophyletic with genus *Triadenum* embedded within (Figs. 2 and 3). *Thornea* is placed in a basal polytomy with the *Elodes–Adenotri*as lineage and the rest of *Hypericum* in the ITS tree (Fig. 2). Phylogenetic relationships within *Hypericum* are also congruent among markers (Figs 2 and 3), showing species from sections *Elodes* and *Adenotrias* (*H. elodes* and *H. aegypticum*) as the sister-group of the remaining species, either forming a clade (A: *Elodes–Adenotrias*) in the cpDNA tree (Fig. 3) or a basal polytomy in the nuclear phylogeny (Fig. 2). Branching next is a sister-group relationship between a mainly New World clade (clade B) and an Old World clade (clades C–E). The New World lineage comprises species belonging to American sections *Myriandra*, *Brathys*, and *Trigynobrathys*, with genus *Triadenum* as their sister-group. The Old World lineage is divided into three major clades C, D, and E, grouping species from Europe, Asia and Africa, but also from Oceania and the New World. Several monophyletic groups or subclades can be recognized within each major clade, which are also geographically structured but do not conform to the current sectional classification. These groups have been given the name of the section with the largest number of species (e.g., ''*Ascyreia*-group'', Fig. 3).

The following sections were recovered as monophyletic in our analysis: *Myriandra*, *Androsaemum*, *Oligostema*, *Webbia*, *Psorophytum*, *Campylopus*, *Bupleuroides*, *Heterophylla*, *Elodes*, *Thasia* and *Inodora*, though the last seven are monotypic. Other sections were represented in the analysis by one specimen (e.g., *Roscyna*) or clade support was low (e.g., *Hirtella*), so monophyly could not be assessed. The remaining sections (e.g., *Trigynobrathys*, *Campylosporus*, *Hypericum*, *Ascyreia*) were inferred to be paraor polyphyletic (see Section 4, Table 3). In a few cases, con-specific specimens were not grouped together such as in species *H. hookerianum*, *H. lancasteri*, *H. empetrifolium* and *H. aethiopicum* in the cpDNA tree (Fig. 3), or *H. lalandii* and *H. synstylum* in the ITS tree (Fig. 2).

Table 2. Sensitivity analysis of different partitioning strategies. Sensitivity analysis to assess the impact of different partitioning strategies on the Bayesian analysis of the ''All-specimens'' concatenate plastid dataset. ''Unpartitioned'': a single substitution model assigned to all sites; ''Partitioned-Uncorrected'': ''by-gene'' partitioned dataset using the default branch length prior (λ = 10; branch length = 0.1). ''Partitioned-Corrected'': by-gene partitioned dataset using the corrected branch length prior (λ = 100; branch length = 0.01). Results for single-gene analyses with (''Corrected'') or without the lambda correction (''Uncorrected'') are also reported. Abbreviations: –ln L: model likelihood (marginal likelihood) estimated as the average of the harmonic mean of the independent runs, following Kass and Raftery, 1995). ML: Results from the maximum likelihood analysis in GARLI. TL-mean: Mean of total tree length estimated over the two independent Bayesian runs. Lambda (λ): branch length prior parameter.

Bayesian analysis	lnL Unpartitioned (TL mean)	lnL Partitioned-Uncorrected (TL mean)	lnL Partitioned-Corrected (TL mean)	lnL ML analysis (TL mean)
''All-specimens''	19330.09 (3.57)	19306.9 (37.94)	19094.65 (2.722)	18929.92 (2.269)
Single genes	lnL Uncorrected (TL mean)		lnL Corrected (TL mean)	lnL ML analysis (TL mean)
ITS	9615.51 (58.489)		9353.4 (4.274)	8910.47 (3.400)
trnL-trnF	5915.90 (30.461)		5803.4 (3.181)	4256.02 (1.138)
trnS-trnG	7607.17 (18.969)		7503.3 (1.871)	6163.75 (1.644)
psbA-trnH	8211.08 (14.059)		7936.4 (2.817)	7668.11 (2.845)

3.4. Ancestral state reconstruction

Fig. 4 shows Bayesian ASR results for seven diagnostic morphological characters. In general, uncertainty was low and most ancestral nodes were reconstructed with posterior estimates over 95%. Our results suggest that the ancestor of *Hypericum* was a darkglandless shrub characterized by three fasciclodes, reticulate seed testa, stellate corolla and three stamen fascicles partially united forming a tube. The herbaceous habit seems to have evolved multiple times in the history of the group, and it is also reconstructed as the ancestral state of the largest clade E (Fig. 4); in contrast, the tree habit is an autapomorphy of the ''Afromontane-group'' in clade D. Dark glands have also evolved independently in clades A, D and E. Othe characters that evolved in parallel in different clades are the pseudo-tubular corollas in clade A and *Triadenum* within clade B, and the presence of five stamen fascicles in clades D and B (with the exception of *Triadenum*, Fig. 4).

3.5. Molecular dating

The crown age of Hypericaceae was estimated at 53.8 Ma with a very broad confidence interval (CI 43 – 66 Ma; SI Appendix). Divergence between tribes Hypericeae (=*Hypericum*) and Vismieae occurred during the Early Eocene (49.9 Ma; CI 41 – 60 Ma), while crown-group *Hypericum* is dated as Late Eocene, 34.9 Ma (CI 34 – 37 Ma). Divergence between the New World and Old World groups is dated in the Eocene–Oligocene boundary (33.7 Ma; CI 30 – 37 Ma), whereas divergence within the three major clades is dated as Early Oligocene (SI Appendix). In general, confidence intervals were small, except for some early divergences, such as the root node, the split of tribe Vismieae, and the crown-age of Clade A.

3.6. Biogeographic analysis

Bayesian ASR of biogeographic ranges in MrBayes was also very decisive (pp > 90), showing Africa as the ancestral area of Hypericaceae, while the remaining ancestral nodes, including crown-group *Hypericum*, are inferred as originating in the Western Palearctic region (Fig. 5). The only exceptions are the MRCA of clade B, which is reconstructed as Nearctic, and the MRCA of clade D, which is inferred as African (Fig. 5). The BEAST BIB reconstruction showed very similar results, but uncertainty was generally higher, which might be attributed to its higher model complexity, with more free parameters than in the standard discrete model used in MrBayes. Hypericaceae is reconstructed as African (marginal probability p = 0.48), but other less supported scenarios included the Western Palearctic region (p = 0.34).

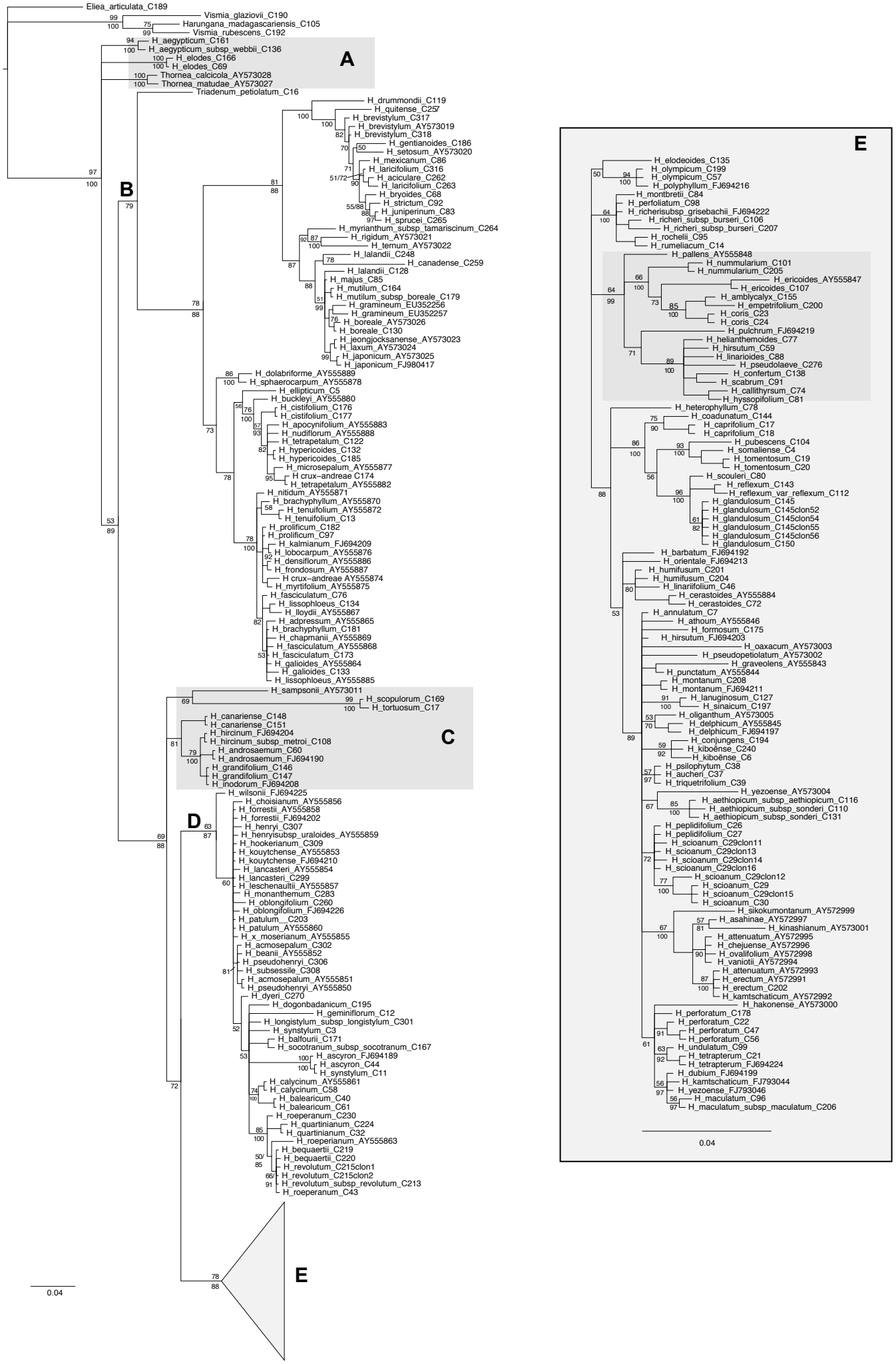


Figure 2. Phylogenetic relationships in Hypericum inferred from ITS sequences. 50% Bayesian Majority-Rule consensus tree with posterior probabilities shown below branches and bootstrap support values for ML rearrangements (500 replicates) above branches. A to E letters indicate major clades discussed in the text. A shaded boxe show a clade that is not well supported by the plastid dataset: "Hirtella-group".

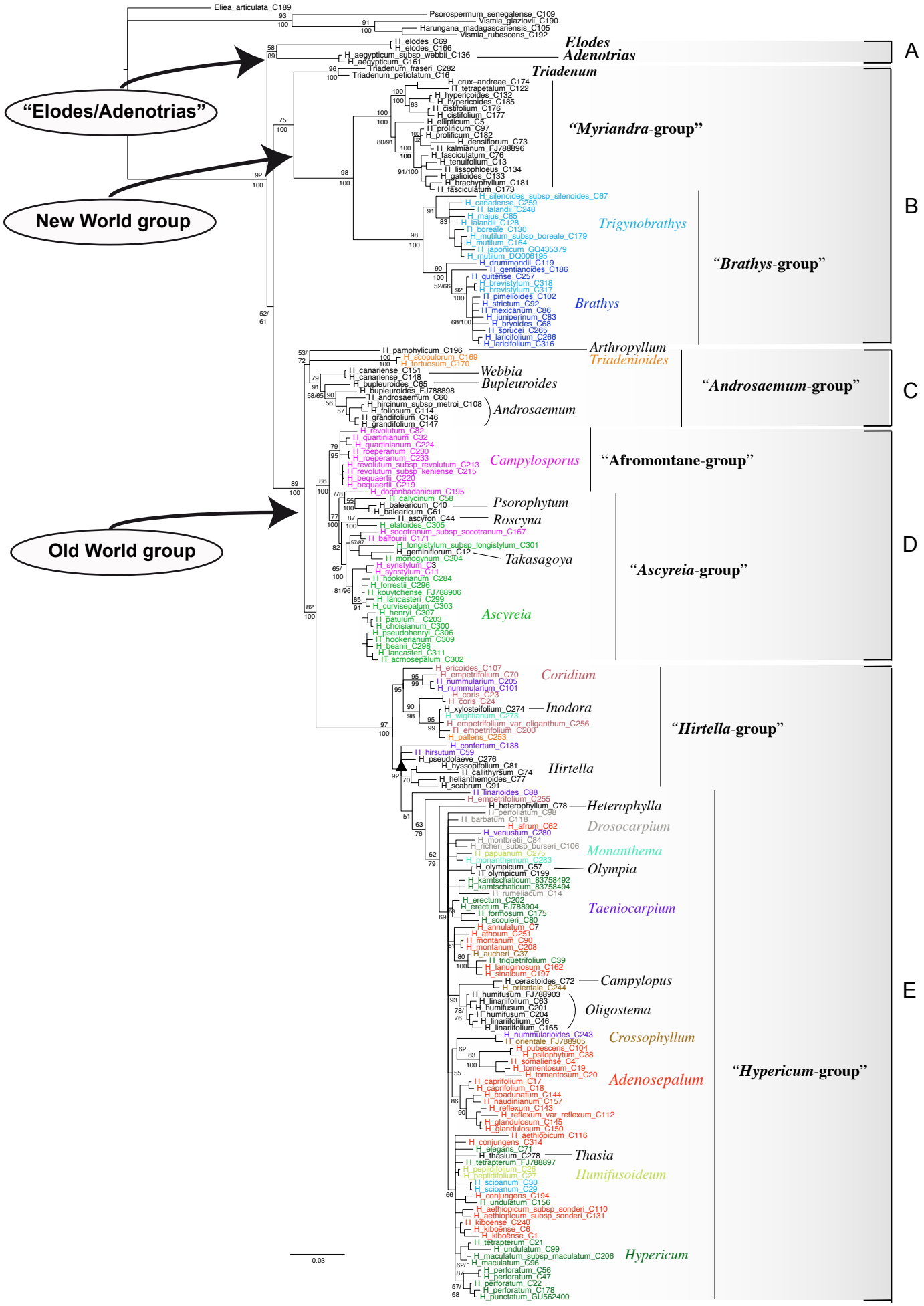


Fig. 3. Phylogenetic relationships in Hypericum inferred from the concatenated "All-specimens" plastid dataset (psbA-trnH, trnL- trnF, trnS-trnG). 50% Bayesian Majority-Rule consensus tree with posterior probabilities shown below branches and bootstrap support values for ML rearrangements (500 replicates) above branches. A black triangle indicates nodes that are not present in the concatenate plastid dataset excluding incomplete taxa ("No-missing"). A to E letters indicate major clades; within them subclades or "groups" are named after the section with the largest number of species included within. Traditional sections (Robson, 1977) discussed in the text are also indicated. Species belonging to sections that were recovered as nonmonophyletic have been highlited by different colours. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

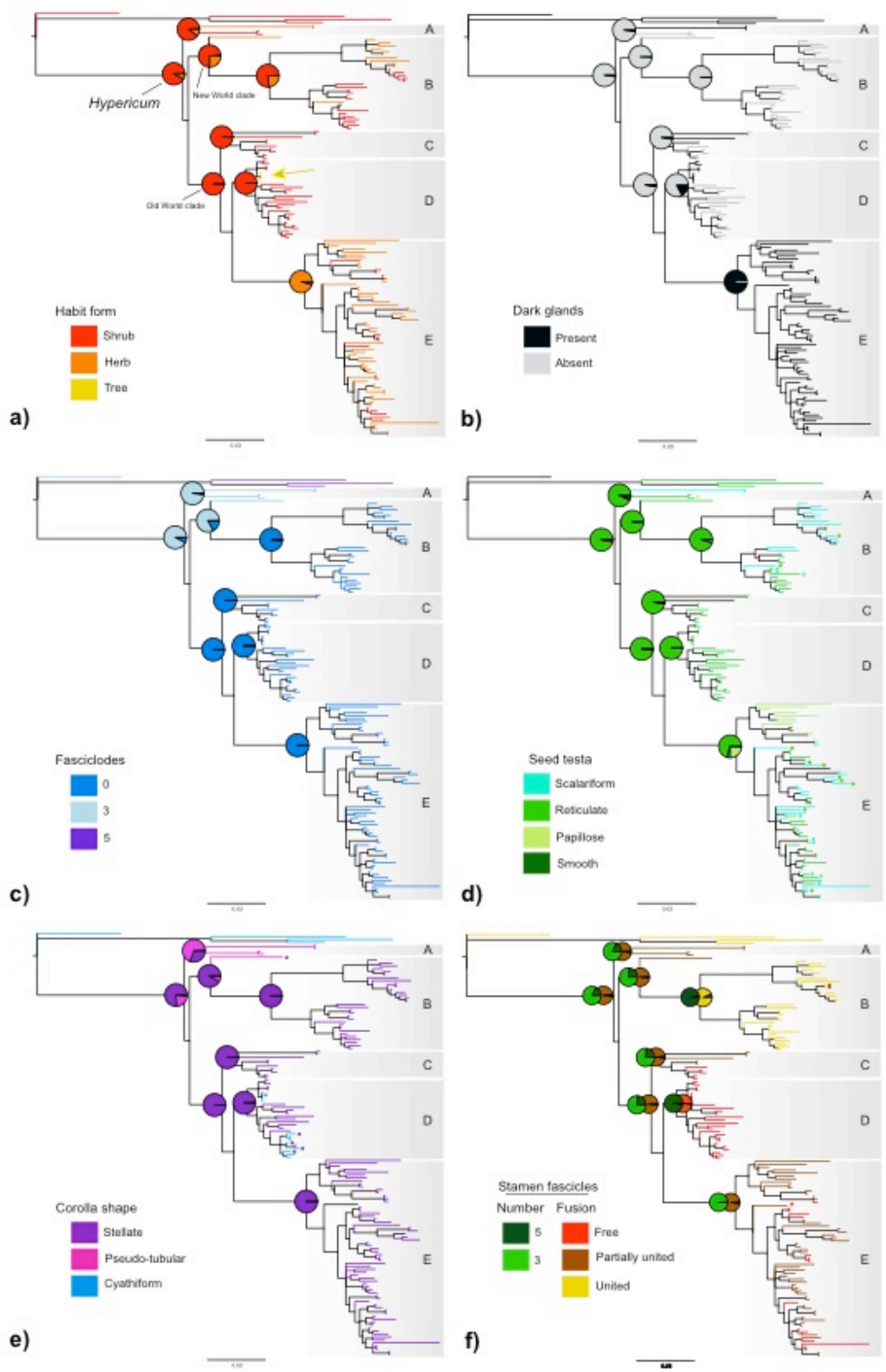


Fig. 4. Bayesian ancestral state reconstruction of diagnostic characters in *Hypericum*. (a) Habit form, (b) presence of dark glands, (c) fasciclodes number, (d) sculpturing pattern of the seed testa, (e) shape of the corolla, (f) number of stamen fascicles and degree of fusion (SI Appendix). The ''Two markers'' chloroplast dataset and MrBayes were used for the reconstruction. Pie charts show the marginal probability for ancestral states at selected nodes, corresponding to the main clades in Fig. 3. Colours on terminal branches represent the character state for each species; black lines indicate missing information (except in b where there was no missing information). Some species were polymorphic (i.e., more than one character state), with one state indicated by the line and the other by a colored dot at the tip. The yellow arrow in (a) highlight the treelet habit within clade D.

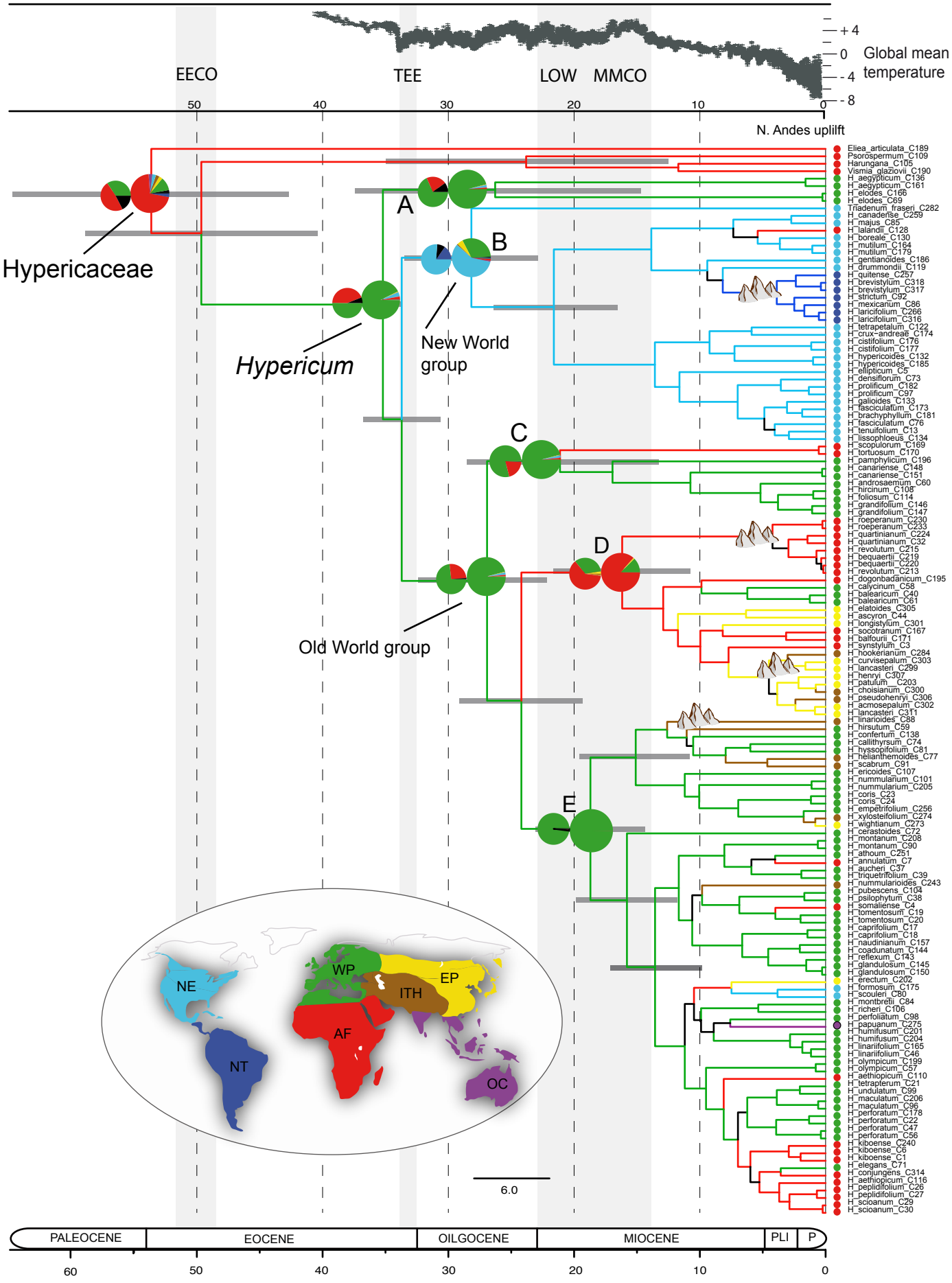


Fig. 5. Bayesians ancestral range reconstruction and molecular dating analysis in Hypericum. Maximum clade credibility (MCC) tree from the BEAST analysis showing median divergence times and 95% confidence intervals (for main lineages) in Hypericum, derived from the ‘‘Two-marker’’ concatenate dataset. Colored branch lengths represent the ancestral range with highest marginal probability for each lineage as inferred with the discrete phylogeographic model of Lemey et al. (2009), implemented in BEAST. Node pie charts represent marginal probabilities for alternative ancestral distributions obtained with MrBayes ancestral state reconstruction (large charts) and BEAST (small charts). Colors correspond with the discrete areas in the inset map. Black lines indicate branches that did not receive clade support. Colored circles before the species name give present ranges. Global mean temperature curve obtained from Zachos et al. (2001); shadow gray vertical bars indicate major climatic events during the Tertiary. Cartoon Mountains show major phases of mountain building. Abbreviations: EECO = Early Eocene Climatic Optima, TEE = Terminal Eocene Event, LOWE = Lower Oligocene Warming Event, MMCO = Mid Miocene Climatic Optimum, Pli = Pliocene, P = Pleistocene. Areas: AF (Sub-Saharian Africa), WP (western Palearctic), EP (eastern Palearctic), IT (IranoTuranian–Himalayan), OC (Oceania), NE (Nearctic), NT (Neotropic).

The area of origin of *Hypericum* is most probably WP (p = 0.5), although Africa was again included among the most likely ancestral areas (p = 0.44). Several dispersal events back to Africa can be observed in the BEAST MCC reconstruction, most notably within Clade C (in the lineage of *H. tortuosum and H. scopulorum*) and in Clade D (e.g., the ''Afromontane-group''). Within Clade D, several dispersal events from Africa to the Western Palearctic, Eastern Palearctic, and Irano-Turanian–Himalayan region are reconstructed. Dispersal from the Nearctic region to the Neotropics is inferred within Clade B, concurrent with the Andean *Brathys* radiation. The most complex migration pattern is found in Clade E, with dispersal events from the Western Palearctic towards Africa, Irano-Turanian–Himalayan region, and the Eastern Palearctic but also trans-oceanic dispersal to the Nearctic and Oceania (Fig. 5). Coding the widespread terminals for the alternative area did not affect biogeographic reconstruction within *Hypericum* i.e., all nodes were reconstructed identically to those in Fig. 5. The only difference was the root of the tree and the ancestor of Hypericeae– Vismieae (crown node Hypericaceae), which were inferred as Western Palearctic, instead of African.

4. Discussion

4.1. Congruence among markers

Our ITS tree was generally congruent with the cpDNA phylogeny, recovering the same major clades and sectional relationships (Figs. 2 and 3). It also agrees well with Nürk et al.'s (2012) ITS phylogeny, showing the early divergent sections ''*Elodes–Adenotrias*'' as sister-group to a geographic dichotomy between a New World clade and an Old World clade. It is difficult to evaluate the congruence at distal levels, since the same taxa were not included in the two studies, and support is generally low for ITS phylogenies (Crockett et al., 2004; Nürk et al., 2012; Park and Kim, 2004; Pilepić et al., 2011). However, we found some cases of well-supported (>95 pp) incongruence between the ITS and the plastid trees in our study that affected low taxonomic levels. Several causes may explain incongruence between gene trees, ranging from hybridization, incomplete lineage sorting, positive selection, paralogy, or poor model choice. The ITS marker may also be affected by problems with homoplasy resulting from extensive sequence variation, compensatory base change, and indel accumulation (Álvarez and Wendel, 2003). Although some of these phenomena may be less relevant at long temporal scale, information from several genetic markers is advisable when inferring the species tree. Chloroplast markers are assumed to not been subject to the same recombination problems as multi-copy nuclear genes. In our case, the concatenated plastid phylogeny also shows better support levels and resolution than the ITS tree, which makes it more appropriate to solve species level relationships. In a multi-gene analysis, overall mutation rates might differ among partitions, and this can cause the overestimation of branch lengths in Bayesian partitioned inference (Brown et al., 2010). We found that this may affect also single-gene analyses when the rate of mutation differs greatly among sites or regions. Correction of the branch length prior helped recovering more realistic branch lengths, comparable with those inferred by ML. Interestingly, Nürk et al. (2012) reported ITS branch lengths that were orders of magnitude longer than our corrected branch lengths (Fig. 2) – but similar to our uncorrected ones (Table 2) which might be explained by their posterior estimate of the phylogeny getting trapped in a region of overly long trees (Brown et al., 2010). Although this has generally no effects on the tree topology (Brown et al., 2010), it might be problematic if branch lengths are later used for inferring lineage divergence times.

Because the plastid genome is haploid and non-recombining, cpDNA markers are expected to show comparable evolutionary histories. Some studies, however, have shown that chloroplast dynamics are sometimes more complicated than assumed, and incongruence between chloroplast genes might reflect underlying biological processes (Medgyesy et al., 1985). Biparental inheritance of cpDNA has been reported in *Hypericum* (Greiner et al., 2011; Renner, 1934). These and other phenomena, such as chloroplast transfer, recombination, or complex mutational dynamics could lead to heteroplasmy (more than one type of organelle DNA within individual cells), which could explain the pattern of incongruence observed between *psbA-trnH* and the other markers (SI Fig. 2). In addition, Borsch and Quandt (Borsch and Quandt, 2009) described a very complex molecular structure including several structural mutations, ancient duplications, and inverted repeat regions in *psbA-trnH*. *psbA-trnH* is the marker in our study with the highest indel mutational rate relative to substitutions, and it exhibits higher levels of saturation than the other cpDNA markers (Table 1). Although we cannot discard the evolutionary processes mentioned above, it is more likely that homoplasy related to

its short size (525 bp if gaps are excluded), high levels of variation, and difficulties in alignment due to its secondary structure, are responsible for the incongruities observed in the *psbA-trnH* gene tree.

4.2. Circumscription of Hypericum

Our phylogenetic results based on plastid and nuclear data are congruent with the division of Hypericaceae into three tribes: Cratoxyleae, Vismieae, and Hypericeae, but reject the monophyly of *Hypericum* (Figs. 2 and 3). Genus *Triadenum* is included within the New World group (clade B) in agreement with previous studies (Nürk et al., 2012; Ruhfel et al., 2011). Nürk et al. (2012) placed Thornea as *Hypericum* sister group whereas Ruhfel et al. (2011) considered this genus as part of *Hypericum*. Our ITS tree places *Thornea* in a polytomy with section *Elodes–Adenotrias* and the rest of *Hypericum*, so we cannot confirm its affiliation. The circumscription of *Hypericum* has long been controversial with different authors including within the Hypericeae genera *Santomasia*, *Lianthus*, *Thornea*, and *Triadenum* (Bentham, 1862; Choisy, 1821; Keller, 1925, 1983), and others excluding the *Hypericum* sections *Elodes* and *Adenotrias* (Kimura, 1951; Spach, 1836a, 1836b). One of the most discussed characters is the presence of fasciclodes between the stamen fascicles. Fasciclodes are absent in the majority of *Hypericum* species, but are present in other tribes and genera of Hypericeae, varying in number from five (tribe Vismieae and genus *Santomasia*) to three (tribe Cratoxyleae, and Hypericeae genera *Lianthus*, *Thornea*, and *Triadenum*). Species from sections *Elodes* and *Adenotrias* are the only ones in *Hypericum* that exhibit (three) fasciclodes. Our Bayesian ASR reconstruction (Fig. 4) based on plastid data agrees with Nürk et al. (2012) in inferring the presence of fasciclodes as ''ancestral'' (plesiomorphic) within *Hypericum*. Other distinctive character is the shape of the corolla, which is stellate in most *Hypericum* species (the ''ancestral'' state, Fig. 4) but pseudotubular (petals are oblique to erect, given the impression of a pseudo-tubular flower) in *Triadenum* and the *Elodes–Adenotrias* clade. The deep bowl-shaped (''cyathyform'') flowers seem to be a specialization of the "*Ascyreia*-group'' and some ''Afromontane'' Campylosporus (Fig. 4), which was interpreted by Robson (Robson, 1981) as a local specialization to mountain climates. The fact that bird pollination has been observed in some of these species (*H. revolutum*: ASM and JJA personal observation, Janeček et al., 2007; Riegert et al., 2011; *H. lanceolatum*: Michenea et al., 2006) seems to confirm the hypothesis that cyathyform flowers evolved as a specialized character in *Hypericum* (Fig. 4).

4.3. Phylogenetic relationships & sectional classification

Our phylogenetic results (Figs. 2 and 3) suggest that the current sectional classification of *Hypericum* needs to be reconsidered, with twelve sections being paraor polyphyletic, eight monotypic and only three confirmed to be monophyletic (Table 3). Our results are in general comparable with those of Nürk et al. (2012) based on ITS, with the exception that we recovered the sections *Campylosporus*, *Coridium* and *Triadenioides* as not monophyletic (see below). Instead, the phylogeny is divided into several geographically segregated clades. Below, we describe these clades and the main morphological traits that support them (as inferred from our ASR analysis, Fig. 4).

(1) The *Elodes–Adenotrias* lineage (Clade A): The monotypic section *Elodes* and section *Adenotrias* (three species, represented here by *H. aegypticum*) form a clade in the chloroplast phylogeny and the BEAST chronogram (Fig. 3, Fig. 5), whose ancestor is characterized by a shrub habit, absence of dark glands, three fasciclodes, reticulate seed testa, pseudo-tubular corolla, and three partially united stamen fascicles. These lineages have sometimes been excluded from *Hypericum* based on their anomalous flower structures (see above), but our results agree with those of Nürk et al. (2012) in placing them as an early-branching lineage, sister-group to the remaining species.

(2) The New World group (Clade B) comprises species from the genus *Triadenum* sister-group of the American sections *Myriandra*, *Brathys* and *Trigynobrathys*. Unlike Pilepić et al. (2011), we recovered *Myriandra* as monophyletic, but inferred *Trigynobrathys* and *Brathys* as polyor paraphyletic. We propose to merge these sections into a larger "*Brathys*-group" following Nürk et al. (2012). The ancestors of this group were probably shrubs with three fasciclodes, reticulate seed testa, stellate corollas and three partially united stamen fascicles, with five united stamen fascicles as an autapomorphy of section Myriandra and the ''Brathys-group'' (Fig. 4).

Table 3 (next page). Taxonomic infra-generic classification of *Hypericum*. ''Traditional section'' refers to Robson's (1977–2012) morphology based sectional classification, with numerical order following the latter study. The other columns compare this classification with results from our and previous phylogenetic studies, based on morphological data (Nürk and Blattner, 2010) or nuclear (ITS) DNA sequences (Nürk et al., 2012). ''Phylogenetic clade'' refers to the major clades described in Fig. 2 and 3 and main text. If plastid and nuclear trees disagree, we indicate both plastid/nuclear results. ''Phylogenetic status'' indicates whether a section was recovered as monophyletic (m), non-monophyletic (p) or monotypic (mt); (?) indicates that the phylogenetic status of the section could not be confirmed because only one representative was sampled or the species falls in a polytomy with taxa from other sections).

	Traditional section	Phylogenetic clade	Phylogenetic status	Nürk and Blattner (2010)	Nürk et al. (2012)
1	Campylosporus (Spach) R. Keller	D	p	m	m
2	Psorophytum (Spach) Nyman	D	mt	mt	mt
3	Ascyreia Choisy	D	p	p	p
4	Takasagoya (Y. Kimura) N. Robson	D	?	p	?
5	Androsaemum (Duhamel) Gordon	C	m	m	m
6	Inodora Stef.	E	mt	mt	mt
6a	Umbraculoides N. Robson	-	-	mt	-
7	Roscyna (Spach) R. Keller	D	?	m	p
8	Bupleuroides Stef.	C	mt	mt	mt
9	Hypericum	E	p	p	p
9a	Concinna N. Robson	-	-	mt	mt
9b	Graveolentia N. Robson	E	?	p	p
9c	Sampsonia N. Robson	C	?	m	m
9d	Elodeoida N. Robson	E	?	p	p
9e	Monanthema N. Robson	E/D	p/?	p	?
10	Olympia (Spach) Nyman	E	?/m	m	m
11	Campylopus Boiss.	E	mt	mt	mt
12	Origanifolia Stef.	-	-	m	m
13	Drosocarpium Spach	E	p	m	p
14	Oligostema (Boiss.) Stef.	E	m/?	p	m
15	Thasia Boiss.	E	mt	mt	-
16	Crossophyllum Spach	E	p	m	?
17	Hirtella Stef.	E	?	p	p
18	Taeniocarpium Jaub. & Spach	E	p	p	p
19	Coridium Spach	E	p/m	m	m
20	Myriandra (Spach) R. Keller	B	m	m	m
21	Webbia (Spach) R. Keller	C	mt	mt	mt
22	Arthrophyllum Jaub. & Spach	C	?	p	m
23	Triadenioides Jaub. & Spach	C, E	p	p	m
24	Heterophylla N. Robson	E	mt	mt	mt
25	Adenotrias (Jaub. & Spach) R. Keller	A	?	m	?
26	Humifusoideum R. Keller	E	?	p	?
27	Adenosepalum Spach	E	p	p	p
28	Elodes (Adans.) W. Koch[a]	A	mt	mt	mt
29	Brathys (Mutis ex L. f.) Choisy	B	p	p	p
30	Trigynobrathys (Y. Kimura) N. Robson	B	p	p	p

The Old World group is the most diversified in terms of number of species and morphological sections and, based in our phylogenetic results, we estimate it contains approximately 270 of the 496 (60%) described species. It is subdivided into three major clades:

(1) Clade C (''*Androsaemum*-group'') comprises species from sections *Bupleuroides*, *Webbia*, *Androsaemum*, *Sampsonia* (only in ITS), *Triadenioides* and *Arthrophyllum*, the last two falling in a polytomy, and receives moderate or low support in the cpDNA and ITS trees (it is also recovered in the BEAST dated tree). We found that *Triadenioides* is polyphyletic, contrary to Nürk et al. (2012) findings that had a reduced sampling of this section. The ancestor of the group is characterized by a shrub habit, absence of dark glands and fasciclodes, reticulate seed testa, stellate flowers, and three partially united stamen fascicles. Free stamen fascicles seem to be apomorphic of section *Androsaemum*.

(2) Clade D is divided into two clades: the ''Afromontane-group'' of section *Campylosporus* and the ''*Ascyreia*-group'', which includes mainly species from the large Asian section *Ascyreia*, but also from *Roscyna*, *Takasagoya*, and the monotypic *Psorophytum* (Fig. 3). Some African species of *Campylosporus*, *H. synstylum, H. balfourii and H. socotranum*, fall within the ''*Ascyreia*-group'', rendering this section polyphyletic contrary to Nürk et al. (2012); this could be explained because we included a larger sampling of this African section in our study. These species differ from the ''Afromontane-group'' in having deciduous petals and stamens and in the absence of dark glands, all characteristics of the ''*Ascyreia*-group'' (Fig. 4). The ancestors of clade D was a darkglandless shrub with reticulate testa, stellate flowers, and five free stamen fascicles, the latter seem to be autapomorphic of this group. The ''Afromontane-group'' shows also several derived characters, such as the tree habit form, presence of dark glands, and cyathiform corollas.

(3) Clade E is the most numerous and variable concerning distribution and morphology. The ancestor of this clade was characterized by the presence of dark glands and herbaceous

habit, absence of fasciclodes, stellate flowers, three partially united stamens, and reticulate seed testa, although there is considerable variation in the last two characters in the current species (Fig. 4). Although resolution within this clade was low, two subclades or groups can be recognized. The ''*Hirtella*-group'' comprises species from sections *Coridium*, *Monanthema*, *Inodora* and *Triadenioides*, as well as *Taeniocarpium* and *Hirtella*. This group, which was also recovered by Nürk et al. (2012) and Crockett et al. (2004), receives moderate support in the ITS tree, the concatenated ''No-missing'' plastid dataset and some of the individual chloroplast trees (SI Fig. 2, it is also recovered by the BEAST tree, Fig. 5), but not in the combined ''All-specimens'' cpDNA tree (Fig. 3). The rest of species and sections are grouped into the ''*Hypericum*-group'', with generally poor internal resolution (Figs. 2 and 3).

4.4. Spatio-temporal evolution in Hypericum

In line with the tennets of Phylogenetic Biogeography (Brundin, 1966; Hennig, 1966), Robson (1981) hypothesized that there was a parallelism between the morphological and geographic evolution of *Hypericum*. He described evolutionary trends for the main diagnostic characters (''morphoclines''), and noted that these morphoclines were generally correlated with distributional trends, defining ''geomorphoclines'' (Robson, 2006). In particular, Robson hypothesized that the genus originated in Africa before the break up of Gondwana, and that the characters exhibited by the Afromontane species (*H. bequarteri* and *H. revolutum*), such as treelet habit and presence of dark glands, were ancestral in the genus. Geographic spread of *Hypericum* from Africa to other continents would have been accompanied by the appearance of derived traits such as the herbaceous habit and the loss of dark glands.

Our BEAST-BIB reconstruction shows a different scenario (Fig. 5). The ancestors of family Hypericaceae are actually reconstructed as African. Coding for the alternative areas for widespread species did not change ASR within *Hypericum*, but it did favor WP as ancestral area for the root and the ancestor of Hypericeae–Vismieae, although Africa was inferred with similar probability (results not shown). With the exception of *Cratoxylum* in SE Asia and *Vismia* widespread in South America and Africa, all other genera in tribes Vismieae and Cratoxyleae are African, so our sampling of outgroups is probably representative of the distribution of the group. Moreover, a more inclusive analysis on the clusioid clade, including representatives of virtually every genera (Ruhfel, 2011), reconstructed Africa as the ancestral area of Hypericaceae and that of the MRCA of Vismieae and Hypericeae. Therefore, it is likely that Africa is the area of origin for Hypericaceae.

The ancestors of *Hypericum* are inferred to have dispersed from Africa to the western part of Europe in the Early Tertiary (Fig. 5), probably using the dispersal route provided by the collision of the African and Iberian Plates in the Paleocene (Meulenkamp and Sissingh, 2003; Rosenbaum et al., 2002). Colonization of the Northern Hemisphere by *Hypericum* stem-lineages seem to have been concurrent with the climate warming that peaked in the Early Eocene Climatic Optima (EECO in Fig. 5; Zachos et al., 2001). At that time, tropical climates characterized higher latitudes, and a uniform vegetation belt, a mixture of deciduous and evergreen plants, the ''boreotropical forest'', covered the Northern landmasses from Asia to Europe and North America (Tiffney, 1985a, 1985b; Wolfe, 1975). *Hypericum* ancestors were probably tropical shrubs, much like related tribes Vismieae and Cratoxyleae, and could have used these favorable tropical conditions to invade the Holarctic.

Crown-group *Hypericum* is reconstructed as having evolved in the West Palearctic region (Fig. 5), with an initial diversification 35 Ma (CI 34–37 Ma, Fig. 5). This range is within the dates inferred by Ruhfel (2011), who estimated the first diversification in *Hypericum* (crown-age) between 30.8 and 37.3 Ma, depending on the position of *Paleoclusia* (see above). The origin of the crown group *Hypericum* seems to coincide with a dramatic drop in global temperatures and increase in seasonality; the Terminal Eocene Event (TEE in Fig. 5; Zachos et al., 2001). This event promoted the selection of cool-adapted boreotropical elements and the expansion of deciduous vegetation at northern latitudes, the ''mixed-mesophytic forest'' (Tiffney, 1985a, 1985b). Some specializations in *Hypericum* such as the change on habit form and the evolution of unspecialized corollas may be related to the adaptation of these ancestral lineages to the new temperate conditions. On the other hand, Davis et al. (2005) reconstructed the ancestors of Hypericaceae as inhabitants of open woodland habitats in tropical latitudes, which could indicate pre-adaptation to more open environments. However, this result needs to be carefully interpreted since the sampling within the family

was very reduced (only *Vismia* and *Hypericum* were included).

Hypericum might have been part of the Mid-Tertiary mixedmesophytic forest, as evidenced by the appearance of *Hypericum* Early–Mid Miocene seeds on relict assemblages of this forest in West Yunnan (China; Zhao et al., 2004). Our hypothesized scenario of a West Palearctic diversification contrasts with the presence of the oldest fossil remains of *Hypericum* in the Late Eocene of West Siberia (Meseguer and Sanmartín, 2012). This suggests that *Hypericum* ancestors were also distributed in the Eastern Palearctic (area ''EP'' in Fig. 5). Bayesian inference of ancestral states does not allow polymorphic (widespread) ancestors, which might be unrealistic for an old group like *Hypericum* that evolved during a time of major geologic changes. However, the Eastern Palearctic is actually poorly represented in *Hypericum*: most lineages within this region, like the ''*Ascyreia*-group'', are restricted to the southern portion (China, Himalaya), whereas the northern part of EP (where *H. antiquum* was found) is now represented by a few widespread species (Robson, 1981). Moreover, our analysis included a good sampling of these EP lineages (e.g., *Roscyna*, *Takasagoya*, *Monanthema*, and *Hypericum*), so our results cannot be attributed to a biased representation of this region. Instead, it is more likely that large-scale extinction in the northern part of the Eastern Palearctic, associated to the Terminal Eocene Event (TEE) and the Late Tertiary climatic fluctuations (Sanmartin et al., 2001), would explain the disagreement between our reconstruction and the fossil record.

The ancestor of the New and Old World lineages is reconstructed to have dispersed from the Palearctic to North America at the end of the Eocene (Fig. 5). At this time, two land corridors connected all northern landmasses: the North Atlantic Land Bridge (NLAB) and the Beringian Land Bridge, BLB (McKenna, 1983; Tiffney, 1985a, 1985b; Wolfe, 1975). Although the general view is that the NALB only persisted until the Early Eocene (McKenna, 1983; Sanmartin et al., 2001; Tiffney, 1985a, 1985b), some authors suggest a longer connection (Donoghue and Moore, 2003; Gronlie, 1979; Wen, 1999). The southern fringes of the Beringian Bridge were probably suitable for cool-tolerant taxa during the Eocene, and this connection is thought to have lasted until the Late Miocene for temperate taxa (Sanmartin et al., 2001). In any event, it is likely that *Hypericum* ancestors used the geographical proximity of North America and Eurasia and the existence of a uniform forest belt, the Eocene boreotropical forest or its successor, the Oligocene mixed-mesophytic forest, to migrate across the northern landmasses. Davis et al. (2002, 2004) also suggested a northern latitude migration to explain the biogeographic history of the pantropical family Malpighiaceae, and similar hypotheses have been proposed for other plant groups (Donoghue and Smith, 2004; Tiffney, 1985a, 1985b; Wen, 1999; Wen and Ickert-Bond, 2009; Wolfe, 1975; Xiang et al., 1998). The trans-Beringian connection seems to have persisted for *Hypericum* until the Late Miocene, as can be observed in the split between *H. erectum* (Eastern Palearctic) and the Nearctic *H. formosum–H. scouleri* (Fig. 5). Another example is *Triadenum*, which has species in eastern North America and northeast Asia, the latter not included in our study. Diversification within the New World group started in the Early Oligocene in North America, with some taxa migrating to Africa probably by long distance dispersal (*H. lalandii*). Dispersal to South America was concurrent with the rising of mountain chains in Central and northern South America in the Late Miocene ca. 12 Ma (Hoorn et al., 2010). Precisely the last peak of mountain building in the Northern Andes at c. 4.5 Ma (Hoorn et al., 2010) coincides with the start of diversification (crown-node) of the South American radiation in the ''*Brathys*-group'' (Fig. 5).

The Old World clade began also diversifying in the Oligocene within the Western Palearctic region (Fig. 5). From there, several dispersal events to the rest of the world are inferred, which are mainly dated after the Mid Miocene Climatic Optimum (MMCO, Fig. 5). Dispersal events back to Africa occurred at different times, but mostly around the Late Oligocene–Early Miocene and the Late Miocene–Pliocene (Fig. 5). The Oligocene–Early Miocene was a warm and humid period, with wide extensions of rainforests from northern Africa to South Africa (Jacobs, 2004; Plana, 2004). This rainforest was fragmented and replaced by a woodland savannah following the aridification process that started in Africa in the Mid Miocene (Coetzee, 1993). This was the result of a combination of factors, the Eastern uplift of the continent, the closure of the Tethys Sea, and the deterioration of global climatic conditions at the end of the Miocene (Zachos et al., 2001). The geographic disjunction between Africa and WP observed in the MRCA of clade C (the lineage of *H. scopulorum–H.tortuosmum* in Socotra and the Mediterranean–Macaronesian clade *H. pamphylicum–H. grandiflorum*, Fig. 5) could be evidence of a formerly widespread African flora fragmented by these climatic events (Sanmartín et al., 2010). Later dispersals to Africa in the Late

Miocene–Pliocene in clade E are concurrent with the Messinian Salinity crisis (c. 7.2 Ma, Krijgsman et al., 1999) and with a period of high tectonic activity (c. 7–8 Ma) that led to the uplift of the Eastern Arc Mountains and the uplands of West Central Africa with the Cameroon volcanic line (Plana, 2004). Indeed, the diversification of the ''Afromontane-group'' in section *Campylosporus* (clade D) is contemporary with the maximum uplift of the Eastern African Rift system in the Pliocene that ended with the formation of the Ethiopian highlands (5–2 Ma, Sepulchre et al., 2006). Dispersal from Africa to Asia by the ancestors of the ''*Ascyreia*-group'' (clade D) in the Late Miocene (Fig. 5) might have been facilitated by the collision of the Arabian plate with Eurasia (c. 16 Ma) and the uplift of the Red Sea margins (13.8 Ma; Goudie, 2005). Another possibility is that the ''*Ascyreia*-group'' in East Asia (China) is a relict assemblage of the Mid-Tertiary mixed-mesophytic forest, as suggested by the findings of Early–Mid Miocene seeds in ths region (Zhao et al., 2004). This further suggests the possibility of a dispersal event in the opposite direction, from Asia to Africa, and of extinction misleading again our reconstruction. The mixed-mesophytic forest went extinct in Europe and western North America following the drastic climate cooling at the end of the Tertiary, but survived in East Asia and eastern North America (Tiffney 1985a, 1985b).

Hypericum colonization and diversification in the Irano-Turanian–Himalayan region (ITH) is dated during the Late Miocene (Fig 5). The paleogeographic history of this region is complex: it was formed by the collision of the Indian and Arabian plates against Eurasia, and the subsequent rise of several mountains ranges. Periods of major uplift in this region seem to coincide with several dispersal events of *Hypericum* lineages to this region: the ''*Hirtellla*-group'' entered the Iranian Plateau (Fig. 5) after the collision of the Arabian and Eurasian plates that resulted in the Late Miocnee uplift of the Zagros Mountains (10 Ma, Sanmartin, 2003). Similarly, some members of the ''*Ascyreia*-group'' colonized the Himalayan Mountains (Fig. 5) coincident with a major orogenic uplift of the Himalayan range, ca. 7–8 Ma (Wang et al., 2009). From our results, it seems possible that the rising of the Neogene mountain ranges (e.g., Northern Andes, Eastern African Mountains, Himalayan mountains) played an important role in the colonization of tropical and subtropical regions in *Hypericum*, where mountain uplift favoured the appearance of new niches for temperate adapted taxa.

References

Aguilar, J.F., Rosselló, J.A., Feliner, G.N., 1999. Molecular evidence for the compilospecies model of reticulate evolution in *Armeria* (Plumbaginaceae). Systematic Biology 48, 735–754.

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Kiado, A. (Ed.), Second International Symposium on Information Theory, Budapest, pp. 267–281.

Álvarez, I., Wendel, J.F., 2003. Ribosomal ITS sequences and plant phylogenetic inference. Molecular Phylogenetics and Evolution 29, 417–434.

Arbuzova, O., 2005. *Hypericum* L. In: Budantsev, L. (Ed.), Iskopaemye tsvetkovye rastenija Rossii i sopredel'nyh gosudarstv [Fossil Flowering Plants of Russia and Adjacent Countries]. Izdatelstvo Nauka Leningradskoe otdnie, 1974-. Leningrad.

Barnes, J., Anderson, L.A., Phillipson, D.J., 2001. St. John's wort (*Hypericum* perforatum L.): a review of its chemistry, pharmacology and clinical properties. Journal of Pharmacy and Pharmacology 53, 583–600.

Bentham, G., 1862. Hypericineae and Guttiferae. In: Bentham, G., Hooker, J.D. (Eds.), Genera Plantarum, vol. 1, London, pp. 163–177.

Borsch, T., Quandt, D., 2009. Mutational dynamics and phylogenetic utility of noncoding chloroplast DNA. Plant Systematics and Evolution 282, 169–199.

Brown, J.M., Hedtke, S.M., Lemmon, A.R., Lemmon, E.M., 2010. When trees grow too long: investigating the causes of highly inaccurate bayesian branch-length estimates. Systematic Biology 59, 145–161.

Brundin, L., 1966. Transantarctic relationships and their significance, as evidenced by chironomid midges: with a monograph of the subfamilies Podonominae and Aphroteniinae and the Austral Heptagyiae. Kungliga Svenska Vetenskapsakademien Handlingar 11, 1–472.

Carine, M.A., Christenhusz, M.J.M., 2010. About this volume: the monograph of *Hypericum* by Norman Robson. Phytotaxa 4, 1–4.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Molecular Biology and Evolution 17, 540–552.

Choisy, J.D., 1821. Prodromus d'une monographic de la famille des Hypericacees, Geneva.

Coetzee, J.A., 1993. African flora since the terminal Jurassic. In: Goldblatt, P. (Ed.), Biological Relationships between Africa and South America. Yale University Press, New Haven, pp. 37–61.

Crockett, S.L., Douglas, A.W., Scheffler, B.E., Khan, I.A., 2004. Genetic profiling of *Hypericum* (St. John's Wort) species by nuclear ribosomal ITS sequence analysis. Planta Medica 70, 929–935.

Davis, C.C., Bell, C.D., Mathews, S., Donoghue, M.J., 2002. Laurasian migration explains Gondwanan disjunctions: evidence from Malpighiaceae. Proceedings of the National Academy of Sciences of the United States of America 99, 6833–6837.

Davis, C.C., Fritsch, P.W., Bell, C.D., Mathews, S., 2004. High-latitude tertiary migrations of an exclusively tropical clade: evidence from Malpighiaceae. International Journal of Plant Sciences, 165.

Davis, C.C., Webb, C.O., Wurdack, K.J., Jaramillo, C.A., Donoghue, M.J., 2005. Explosive radiation of malpighiales supports a mid-Cretaceous origin of modern tropical rain forests. American Naturalist 165, E36–E65.

Donoghue, M.J., Moore, B.R., 2003. Toward an integrative historical biogeography.Integrative and Comparative Biology 43, 261–270.

Donoghue, M.J., Smith, S.A., 2004. Patterns in the assembly of temperate forests around the Northern Hemisphere. Philosophical Transactions of the Royal Society of London B Biological Sciences 359, 1633–1644.

Doyle, J.J., 1992. Gene trees and species trees: molecular systematics as one-character taxonomy. Systematic Botany 17, 144–163.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. Bmc Evolutionary Biology 7, 8.

Goudie, A.S., 2005. The drainage of Africa since the Cretaceous. Geomorphology 67, 437–456.

Greiner, S., Rauwolf, U., Meurer, J., Herrmann, R.G., 2011. The role of plastids in plant speciation. Molecular Ecology 20, 671–691.

Gronlie, G., 1979. Tertiary paleogeography of the Norwegian Greenland Sea. Norsk Polarinstitutt Skrifter 170, 49–61.

Gustafsson, C., Persson, C., 2002. Phylogenetic relationships among species of the neotropical genus *Randia* (Rubiaceae, gardenieae) inferred from molecular and morphological data. Taxon 51, 661–674.

Hamilton, M.B., 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. Molecular Ecology 8, 521–523.

Heenan, P.B., 2008. Three newly recognised species of *Hypericum* (Clusiaceae) from New Zealand. New Zealand Journal of Botany 46, 547–558.

Hennig, W., 1966. Phylogelletic Systelllotics. University of Illinois Press, Urbana.

Ho, S.Y.W., Phillips, M.J., 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Systematic Biology 58, 367–380.

Hoorn, C., Wesselingh, F.P., Ter Steege, H., Bermudez, M.A., Mora, A., Sevink, J., et al., 2010. Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. Science 330, 927–931.

Huelsenbeck, J.P., Bollback, J.P., 2001. Empirical and hierarchical Bayesian estimation of ancestral states. Systematic Biology 50, 351–366.

Jacobs, B.F., 2004. Palaeobotanical studies from tropical Africa: relevance to the evolution of forest, woodland and savannah biomes. Philosophical Transactions of the Royal Society B: Biological Sciences 359, 1573–1583.

Janeček, Š., Hrázský, Z., Bartoš, M., Brom, J., Reif, J., Hořák, D., et al., 2007. Importance of big pollinators for the reproduction of two Hypericum species in Cameroon, West Africa. African Journal of Ecology 45, 607–613.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. Journal of the American Statistical Association 90, 773–795.

Katoh, K., Toh, H., 2008. Recent developments in the MAFFT multiple sequence alignment program. Briefings in Bioinformatics 9, 286–298.

Katoh, K., Misawa, K., Kuma, K.I., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Research 30, 3059–3066.

Kay, K.M., Whittall, J.B., Hodges, S.A., 2006. A survey of nuclear ribosomal internal transcribed spacer substitution rates across angiosperms: an approximate molecular clock with life history effects. BMC Evolutionary Biology 6, 36. http://dx.doi.org/10.1186/1471-2148-6-36.

Keller, R., 1925. *Hypericum*. In: Engler, A., Prantl, K. (Eds.), Die natürlichen Pflanzenfamilien. Engelmann, Leipzig, pp. 175–183.

Keller, R., 1983. *Hypericum*. In: Engler, A., Prantl, K. (Eds.), Die natürlichen Pflanzenfamilien Leipzig, pp. 208–215.

Kibbe, W.A., 2007. OligoCalc: an online oligonucleotide properties calculator. Nucleic Acids Research 35, W43–W46.

Kimura, Y., 1951. Hypericaceae. In: Nakai, T., Honda, M. (Eds.), Nova Flora Japonica, Tokyo, Sanseido.

Krijgsman, W., Hilgen, F.J., Raffi, I., Sierro, F.J., Wilson, D.S., 1999. Chronology, causes and progression of the Messinian salinity crisis. Nature 400, 652–655.

Lemey, P., Rambaut, A., Drummond, A.J., Suchard, M.A., 2009. Bayesian phylogeography finds its roots. PLoS Computational Biology 5.

Lemmon, A.R., Brown, J.M., Stanger-Hall, K., Lemmon, E.M., 2009. The effect of missing data on phylogenetic estimates obtained by maximum likelihood and Bayesian interference. Systematic Biology 58, 130–145.

Lewis, P.O., 2001. A likelihood approach to estimating phylogeny from discrete morphological data. Systematic Biology 50, 921–925.

Litsios, G., Salamin, N., 2012. Effects of phylogenetic signal on ancestral state reconstruction. Systematic Biology 61, 533–538.

Marshall, D.C., 2010. Cryptic failure of partitioned bayesian phylogenetic analyses: lost in the land of long trees. Systematic Biology 59, 108–117.

Marshall, D.C., Simon, C., Buckley, T.R., 2006. Accurate branch length estimation in partitioned Bayesian analyses requires accommodation of among-partition rate variation and attention to branch length priors. Systematic Biology 55, 993–1003.

McKenna, M.C., 1983. Cenozoic paleogeography of North Atlantic land bridges. Structure and Development of the Greenland–Scotland Ridge, pp. 351–399.

Medgyesy, P., Fejes, E., Maliga, P., 1985. Interspecific chloroplast recombination in a Nicotiana somatic hybrid. Proceedings of the National Academy of Sciences 82, 6960–6964.

Meseguer, A.S., Sanmartín, I., 2012. Paleobiology of the genus *Hypericum* (Hypericaceae): a survey of the fossil record and its palaeogeographic implications. Anales del Jardin Botanico de Madrid 69, 97–106.

Meulenkamp, J.E., Sissingh, W., 2003. Tertiary palaeogeography and tectonostratigraphic evolution of the Northern and Southern Peri-Tethys platforms and the intermediate domains of the African–Eurasian convergent plate boundary zone. Palaeogeography, Palaeoclimatology, Palaeoecology 196, 209–228.

Michenea, C., Jacques, F., Thierry, P., 2006. Bird pollination in an angraecoid orchid on Reunion Island (Mascarene Archipelago, Indian Ocean). Annals of Botany 97, 965–974.

Müller, K., 2005. SeqState: primer design and sequence statistics for phylogenetic DNA datasets. Applied Bioinformatics 4, 65–69.

Nürk, N.M., Blattner, F.R., 2010. Cladistic analysis of morphological characters in *Hypericum* (Hypericaceae). Taxon 59, 1495–1507.

Nürk, N.M., Madriñán, S., Carine, M.A., Chase, M.W., Blattner, F.R., 2012. Molecular phylogenetics and morphological evolution of St. John's wort (*Hypericum*; Hypericaceae). Molecular Phylogenetics and Evolution.

Nylander, J.A.A., 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Nylander, J.A.A., Ronquist, F., Huelsenbeck, J.P., Nieves-Aldrey, J.L., 2004. Bayesian phylogenetic analysis of combined data. Systematic Biology 53, 47–67.

Nylander, J.A.A., Wilgenbusch, J.C., Warren, D.L., Swofford, D.L., 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics 24, 581–583.

Park, S., Kim, K., 2004. Molecular phylogeny of the genus *Hypericum* (Hypericaceae) from Korea and Japan: evidence from nuclear rDNA ITS sequence data. Journal of Plant Biology 47, 366–374.

Pilepić , K.H., Balić , M., Blažina, N., 2011. Estimation of phylogenetic relationships among some *Hypericum* (Hypericaceae) species using internal transcribed spacer sequences. Plant Biosystems 145, 81–87.

Plana, V., 2004. Mechanisms and tempo of evolution in the African Guineo-Congolian rainforest. Philosophical Transactions of the Royal Society B: Biological Sciences 359, 1585–1594.

Rambaut, A., 2002. Se-Al: Sequence Alignment Editor.

Rambaut, A., Charleston, M., 2001. TreeEdit: Phylogenetic Tree Editor v. 1.0 alpha 8. University of Oxford.

Rambaut, A., Drummond, A.J., 2009. Tracer, version 1.5, MCMC Trace Analysis Package.

Ree, R.H., Sanmartín, I., 2009. Prospects and challenges for parametric models in historical biogeographical inference. Journal of Biogeography 36, 1211–1220.

Renner, O., 1934. Die pflanzlichen Plastiden als selbständige Elemente der genetischen Konstitution. Berichte der mathemathisch-physikalischen Klasse der sächsischen Akademie der Wissenschaften zu Leibzig 86, 241–266.

Riegert, J., Fainová, D., Antczak, M., Sedláček, O., Hořák, D., Reif, J., et al., 2011. Food niche differentiation in two syntopic sunbird species: a case study from the Cameroon Mountains. Journal of Ornithology 152, 819–825.

Robson, N.K.B., 1977. Studies in the genus *Hypericum* L. (Guttiferae). 1. Infrageneric classification. Bulletin of the British Museum (Natural History), Botany Series 5, 295-355.

Robson, N.K.B., 1981. Studies in the genus *Hypericum* L. (Guttiferae). 2. Characters of the genus. Bulletin of the British Museum (Natural History), Botany Series 8, 55–226.

Robson, N.K.B., 1985. Studies in the genus *Hypericum* L. (Guttiferae). 3. Sections 1. *Campylosporus* to 6a. *Umbraculoides*. Bulletin of the British Museum (Natural History), Botany Series 12, 163–211.

Robson, N.K.B., 1987. Studies in the genus Hypericum L. (Guttiferae). 7. Section 29. Brathys (part 1). Bulletin of the British Museum (Natural History), Botany Series 16, 1–106.

Robson, N.K.B., 1990. Studies in the genus Hypericum L. (Guttiferae). 8. Sections 29. Brathys (part 2) and 30. Trigynobrathys. Bulletin of the British Museum (Natural History), Botany Series 20, 1–151.

Robson, N.K.B., 1996. Studies in the genus *Hypericum* L. (Guttiferae). 6. Sections 20. *Myriandra* to 28. *Elodes*. Bulletin of the British Museum (Natural History), Botany Series 26, 75–217.

Robson, N.K.B., 2001. Studies in the genus *Hypericum* L. (Guttiferae). 4(1). Sections 7. *Roscyna* to 9. *Hypericum* sensu lato (part 1). Bulletin of the British Museum (Natural History), Botany Series 31, 37–88.

Robson, N.K.B., 2002. Studies in the genus *Hypericum* L. (Guttiferae).4(2). Section 9. *Hypericum* sensu lato (part 2): subsection 1. *Hypericum* series 1, *Hypericum*. Bulletin of the Natural History Museum, London (Botany) 32, 61–123.

Robson, N.K.B., 2006. Studies in the genus *Hypericum* L. (Clusiaceae). Section 9. *Hypericum* sensu lato (part 3): subsection 1. *Hypericum* series 2. *Senanensia*, subsection 2. *Erecta* and section 9b *Graveolentia*. Systematics and Biodiversity 4, 19–98.

Robson, N.K.B., 2010a. Studies in the genus *Hypericum* L. (Hypericaceae). 5(1). Sections 10. *Olympia* to 15/16. *Crossophyllum*. Phytotaxa 4, 5–126.

Robson, N.K.B., 2010b. Studies in the genus *Hypericum* L. (Hypericaceae). 5(2). Section 17. *Hirtella* to 19. *Coridium*. Phytotaxa 4, 127–258.

Robson, N.K.B., 2012. Studies in the genus *Hypericum* L. (Hypericaceae) 9. Addenda, corrigenda, keys, lists and general discussion. Phytotaxa 72, 1–111.

Ronquist, F., 2004. Bayesian inference of character evolution. Trends in Ecology and Evolution 19, 475–481.

Ronquist, F., Sanmartín, I., 2011. Phylogenetic methods in historical biogeography. Annual Review of Ecology, Evolution, and Systematics.

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Hohna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. Mrbayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61, 539–542.

Rosenbaum, G., Lister, G.S., Duboz, C., 2002. Reconstruction of the tectonic evolution of the western Mediterranean since the Oligocene. Journal of the Virtual Explorer 8, 107–130.

Ruhfel, B.R., 2011. Systematics and Biogeography of the Clusioid Clade (Malpighiales). Harvard University, Cambridge, Massachusetts.

Ruhfel, B.R., Bittrich, V., Bove, C.P., Gustafsson, M.H.G., Philbrick, C.T., Rutishauser, R., et al., 2011. Phylogeny of the clusioid clade (Malpighiales): evidence from the plastid and mitochondrial genomes. American Journal of Botany 98, 306–325.

Sanderson, M.J., 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Molecular Biology and Evolution 14, 1218–1231.

Sanmartin, I., 2003. Dispersal vs. vicariance in the Mediterranean: historical biogeography of the Palearctic Pachydeminae (Coleoptera, Scarabaeoidea). Journal of Biogeography 30, 1883–1897.

Sanmartin, I., Enghoff, H., Ronquist, F., 2001. Patterns of animal dispersal, vicariance and diversification in the Holarctic. Biological Journal of the Linnean Society 73, 345–390.

Sanmartin, I., van der Mark, P., Ronquist, F., 2008. Inferring dispersal: a Bayesian approach to phylogeny-based island biogeography, with special reference to the Canary Islands. Journal of Biogeography 35, 428–449.

Sanmartín, I., Anderson, C.L., Alarcon, M., Ronquist, F., Aldasoro, J.J., 2010. Bayesian island biogeography in a continental setting: the Rand Flora case. Biology Letters 6, 703–707.

Sepulchre, P., Ramstein, G., Fluteau, F., Schuster, M., Tiercelin, J.J., Brunet, M., 2006. Tectonic uplift and Eastern Africa aridification. Science 313, 1419–1423.

Simmons, M.P., 2012. Misleading results of likelihood-based phylogenetic analysesin the presence of

missing data. Cladistics 28, 208–222.
Simmons, M.P., Ochoterena, H., 2000. Gaps as characters in sequence-based phylogenetic analyses. Systematic Biology 49, 369–381.
Smith, S.A., Dunn, C.W., 2008. Phyutility: a phyloinformatics tool for trees, alignments and molecular data. Bioinformatics 24, 715–716.
Spach, E., 1836a. Conspectus monographiae Hypericacearum. Annales des Sciences Naturelles – Botanique et Biologie Vegetale 5, 349–369.
Spach, E., 1836b. Hypericacearum monographiae fragmenta. Annales des Sciences Naturelles – Botanique et Biologie Vegetale 5, 156–176.
Stamatakis, A., Hoover, P., Rougemont, J., 2008. A rapid bootstrap algorithm for the RAxML web servers. Systematic Biology 57, 758–771.
Stevens, P.F., 2007. Hypericaceae. In: Kubitzki, K. (Ed.), The Families and Genera of Vascular Plants. Springer, Berlin, Heidelberg, pp. 194–201.
Swofford, D.L., 2002. PAUP : Phylogenetic Analysis Using Parsimony (and Other Methods). Version.
Taberlet, P., Gielly, L., Pautou, G., Bouvet, J., 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Molecular Biology 17, 1105–1109.
Thiers, B., 2008. Index herbariorum: a global directory of public herbaria and associated staff. New York Botanical Garden. http://sweetgum.nybg.org/ih/.
Tiffney, B.H., 1985a. Perspectives on the origin of the floristic similarity between eastern Asia and eastern North America. Journal of the Arnold Arboretum 66, 73–94.
Tiffney, B.H., 1985b. The Eocene North Atlantic land bridge: its importance in Tertiary and modern phytogeography of the Northern Hemisphere. Journal of the Arnold Arboretum 66, 243–273.
Walker, J.D., Geissman, J.W., 2009. 2009 GSA geologic time scale. GSA Today 19, 60.
Wang, Y.J., Susanna, A., Von Raab-Straube, E., Milne, R., Liu, J.Q., 2009. Island-like radiation of Saussurea (Asteraceae: Cardueae) triggered by uplifts of the Qinghai–Tibetan Plateau. Biological Journal of the Linnean Society 97, 893–903.
Wen, J., 1999. Evolution of eastern Asian and eastern North American disjunct distributions in flowering plants. Annual Review of Ecology and Systematics 30, 421–455.
Wen, J., Ickert-Bond, S.M., 2009. Evolution of the Madrean–Tethyan disjunctions and the North and South American amphitropical disjunctions in plants. Journal of Systematics and Evolution 47, 331–348.
White, T.J., Bruns, T., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J. (Eds.), PCR protocols: a guide to methods and applications. Academic Press, San Diego, pp. 315–322.
Wiens, J.J., 2006. Missing data and the design of phylogenetic analyses. Journal of Biomedical Informatics 39, 34–42.
Wiens, J.J., Morrill, M.C., 2011. Missing data in phylogenetic analysis: reconciling results from simulations and empirical data. Systematic Biology 60, 719–731.
Wolfe, J.A., 1975. Some aspects of plant geography of the northern hemisphere during the Late cretaceous and Tertiary. Annals of the Missouri Botanical Garden 62, 264–279.
Wurdack, K.J., Davis, C.C., 2009. Malpighiales phylogenetics: gaining ground on one of the most recalcitrant clades in the angiosperm tree of life. American Journal of Botany 96, 1551–1570.
Xiang, Q.Y., Soltis, D.E., Soltis, P.S., 1998. The eastern Asian and eastern and western north American floristic disjunction: congruent phylogenetic patterns in seven diverse genera. Molecular Phylogenetics and Evolution 10, 178–190.
Zachos, J., Pagani, M., Sloan, L., Thomas, E., Billups, K., 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. Science 292, 686–693.
Zhao, L.C., Wang, Y.F., Liu, C.J., Li, C.S., 2004. Climatic implications of fruit and seed assemblage from Miocene of Yunnan, southwestern China. Quaternary International 117, 81–89.
Zwickl, D.J., 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence dataset under the maximum likelihood criterion. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion.

CHAPTER 2

Utility of low-copy nuclear markers in phylogenetic reconstruction of *Hypericum* L. (Hypericaceae)

This chapter has been done in collaboration with Isabel Sanmartín from the Department of Biodiversity and Conservation, Real Jardín Bótanico-CSIC, Spain, and Thomas Marcussen and Bernard Pfeil, from the Department of Biological and Environmental Sciences, University of Gothenburg (Sweden)

This chapter has been submitted to the peer-review journal *Plant Systematics and Evolution*, and currently it is under review.

CHAPTER 2

Utility of low-copy nuclear markers in phylogenetic reconstruction of *Hypericum* L. (Hypericaceae)

Andrea Sanchez Meseguer, Isabel Sanmartín, Thomas Marcussen, Bernard Pfeil

A b s t r a c t

Primers and sequence variation for two low-copy nuclear genes (LCG) not previously used for phylogenetic inference in the genus *Hypericum*, PHYC and EMB2765, are presented here in comparison with the fast-evolving nuclear intergenic spacer ITS. Substitution rates in the LCG markers were half those reported in ITS for *Hypericum*, which might help avoid the problems caused by substitution saturation and difficulties to establish homologies that afflict the latter marker. Levels of phylogenetic resolution, clade support values and internal character consistency were similar to, or even higher than, those of ITS-based phylogenies. We found evidence for the presence of at least two copies in EMB2765 in *Hypericum*. This methodological challenge was circumvented by the design of an effective clade-specific primer. Both EMB2765 but especially PHYC appear to be good alternatives to the ITS marker, confirming the main phylogenetic relationships found in previous studies, but withimproved resolution and support values for some basal relationships.

1. Introduction

In biosystematics, species phylogenies are generally estimated from gene phylogenies. As the gene phylogenies are contained within the species phylogeny, they represent different levels of organisation and will differ in both topology and relative branch lengths. Processes at both levels can produce gene-to-gene inconsistences or gene-tree/species tree conflicts that may hinder the reconstruction of the species phylogeny (Doyle 1992). Processes acting at the gene level typically affect internal branch lengths and branching order, e.g., gene duplication/extinction and incomplete lineage sorting (ILS). Processes at the species level typically result in inter-species reticulations, e.g., by introgression, homoploid hybrid speciation, or allopolyploid speciation. Thus, species phylogenies are often networks that cannot be fully represented by a branching tree model (Legendre and Makarenkov 2002).

The recognition of the difference between gene phylogenies and the species phylogeny has led to advice in favour of using multiple unlinked data sets (Small et al. 2004). In phylogenetic analysis, multiple markers are sometimes analysed jointly by concatenating genes into a single dataset without assessing for incongruence (Gatesy et al. 1999).

However, this approach is problematic in the sense that it ignores incongruence among individual gene phylogenies. This can in some cases obscure the phylogenetic signal (Pfeil 2009) and/or produce spurious relationships, such as in the case of ILS (Kubatko et al. 2007), hybridisation (Ballard 2000), and putative positive selection (Stefanovic, et al. 2009). Instead, new methods have been developed to deal with gene-to-gene inconsistencies, discriminate between different types of gene incongruence, and more powerfully estimate the species phylogeny, in what is nowadays an active field of research (Meng and Kubatko 2009; Heled and Drummond 2010; Maureira-Butler et al. 2008; Joly et al. 2009; Frajman et al. 2009; Bloomquist et al. 2010; Blanco-Pastor et al. 2012; Holland et al. 2008; Ané et al. 2007; Yu et al. 2011, Jones et al. 2013).

All these methods have in common that they require the use of several independent loci or markers for disentangling the role of alternative biological processes and recovering the species phylogeny.

The nuclear ribosomal inter-transcribed spacer ITS is by far the most widely used marker in plant (as well as fungal) systematics because of the facility with which it can be amplified using near-universal primers (White et al. 1990; Baldwin et al. 1995). However, its particular structure, with large tandem arrays of hundreds or thousands of more or less similar copies of which the amplified sequence is a weighted average, can lead to problems of incomplete concerted evolution, associated paralogy issues, and the presence of non-functional pseudogene copy types (Wendel et al. 1995; Alvarez and Wendel 2003; Nieto-Feliner and Roselló 2007). Lower thermodynamical stability and a higher rate of mutation than other ribosomal markers can also cause problems in phylogenetic inference (Mayol and Rosello 2001) or estimation of lineage divergence times (Kay et al. 2006).

Low-copy or single copy nuclear genes (LCGs) represent a vast but generally unexplored number of unlinked genetic markers rich in phylogenetic information (Sang 2002). LCGs have proven useful to resolve relationships at low taxonomic levels and, unlike the fast evolving nuclear ITS, are expected to exist in one copy per chromosome set (Crawford and Mort 2004; Small et al. 2004), an assumption that is readily testable as more genomes are sequenced. Nevertheless, nuclear gene families may have complex evolutionary dynamics. For example, duplicate gene copies are frequently found in plant genomes as a consequence of both local (paralogues) and genome-wide (homologues) duplication processes (Clegg et al. 1997; Innes et al. 2008). The occurrence of paralogues or homologues creates phylogenetic methodological difficulties, which has made more challenging the extensive use of LCGs in plant systematics (Sang 2002). In the ideal case, a set of primers would amplify a gene at a single locus, making downstream analysis fairly straightforward (allelic variation notwithstanding). However, primers for a gene that is single copy in one clade or pilot study (e.g., Denton et al 1998) may instead amplify more than one copy in another clade (e.g., Oxelman et al. 2004; Pfeil et al. 2004; Ekenäs et al. 2012), thus making the transfer of existing primers to new groups challenging.

Hypericum L. is the largest genus within the family Hypericaceae. It comprises nearly 500 species of shrubs, small trees or rhizomatous, sometimes annual, herbs, with yellow flowers and frequently glandulous tepals or leaves (Robson 2012). The present diversity of the genus has been classified in 36 morphological sections distributed worldwide and covering different environments – *Hypericum* is only absent from the poles, deserts and low-altitude tropical areas (Robson 1981). The most recent systematic revisions included the family Hypericaceae within the informal "clusioid clade" of order Malpighiales, which also includes tropical families such as Clusiaceae, Bonnetiaceae, and Podostemaceae (Davis et al. 2005; Rufhel et al. 2011). Hypericaceae comprises three tribes: the mainly tropical tribes Vismieae (*Vismia*, *Harungana*, and *Psorospermum*) and Cratoxyleae (*Cratoxylum*, *Eliea*), and tribe Hypericeae, which includes the genera *Triadenum, Thornea, Santomasia, Lianthus*, and *Hypericum* (Rufhel et al. 2011; Stevens 2007). Recent molecular work has shown that *Hypericum* is paraphyletic to *Triadenum, Thornea*, and *Santomasia* (Ruhfel et al. 2011; Nürk et al. 2013; Meseguer et al. 2013). These studies also reject the traditional infrageneric classification, recovering many of the large taxonomic sections as non-monophyletic (e.g., *Ascyreia*, *Hirtella*, *Hypericum*, and *Brathys*). To date, all molecular studies have relied on either chloroplast markers or the nuclear ribosomal ITS marker. Meseguer et al. (2013) compared the phylogenetic signal of these two genomes and found overall congruence, supporting a geographical dichotomy between a New World group, comprising the sections *Myriandra*, *Brathys*, and *Trigynobrathys* and the genus *Triadenum*, and an Old World group, comprising the remaining species and sections, e.g. *Ascyreia*, *Hypericum, Campylosporus*, and *Hirtella*. The Western Palearctic, species-poor sections *Elodes* and *Adenotrias* form the sister-group to the New World-Old World clade, although this relation needs to be clarified as it received little support. Meseguer et al. (2013) also reported some cases of incongruence between nuclear and plastid markers, mainly affecting species or species groups, and a general lack of support for both basal and distal relationships in the ITS phylogeny.

Numerous studies suggest that LCGs have the potential to compensate for the lack of resolution and support values of phylogenies based on cpDNA and nrDNA, as well as the ability to recover reticulate phylogenetic relationships (Sang 2002; but see Rauscher et al. 2002). Previous efforts to use LCGs in *Hypericum* have been limited to two species of pharmacological importance, *H. perforatum* and *H. androsaemum*, to study protein expression in relation to hypericine biosynthesis and the genes encoding it (Liu et al., 2003; Bais et al. 2003; Karppinen and Hohtola 2008). Wurdack and Davis (2009) explored the use of LCGs in resolving phylogenetic relationships within order Malpighiales, and concluded that some of these genes, in particular the rapidly evolving PHYC and EMB2765, could be useful in the systematics of this order. Here, we evaluate the utility of LCGs in resolving relationships in genus *Hypericum*, and between *Hypericum* and its closest relatives, and explore their potential to improve branch support values and resolution in comparison with the nuclear ribosomal ITS marker (Meseguer et al. 2013). We assess levels of variation for the two low copy nuclear regions PHYC and EMB2765, and present newly developed PHYC primers specific to *Hypericum* and clade-specific primers to isolate paralogous copies in EMB2765.

2. Materials and methods

2.1 Taxonomic and gene sampling

Species sampling included representatives of 13 out of 36 morphological sections of *Hypericum* and focused on representing all major clades within the group as found in a comprehensive phylogenetic analysis of the genus based on ITS and three different plastid markers (*trnL-trnF*, *trnS-trnG*, *psbA-trnH*; Meseguer et al. 2013). The sample also included representatives of closely related genera and families. DNA was extracted from fresh material collected in the field and preserved in silica gel, and from dry material preserved at several herbaria. GenBank accessions from previous studies were also included (Table 1).

2.2 Amplification and Sequencing

We initially screened eight low-copy nuclear regions, using primers published in the literature and others newly designed in this study for *Hypericum* (Table 2), PCR products range from 800 to 1100 base pairs. The regions were: phytochromeC (PHYC), embryo-defective 2765 (EMB2765, "At2g38770"), chalcone synthase (CHS), waxy (GSSBI), chloroplast-expressed glutamine synthase (ncpGS, GS2, glnII or gln), glucose-6-phosphate isomerase (GPI, "PGIc"), salt tolerance during germination 1 (STG1, TAFII15, "At4 g31720"), and beta-carotene hydroxylase (Chyb). Internal primers designed by Wurdack and Davis (2009) in Malpighiales were initially used to amplify PHYC and EMB2765 markers, but later we designed a new set of PHYC primers specific to *Hypericum* to increase the length of the amplified region. For EMB2765, we also designed clade-specific primers to isolate paralogous copies. For comparative purposes, we also sequenced the ITS region using primers ITS1 and ITS4 (Aguilar et al. 1999; White et al. 1990; Table 1). DNA was extracted from leaf tissue samples using the QIAGEN DNeasy plant kit (Qiagen, Hilden, Germany) at the laboratories of the Real Jardín Botánico-CSIC (Madrid, Spain), and following the manufacturer's protocol. The PCR cycling conditions were as follows: 95ºC for 5 min, 35 cycles of [94ºC for 30 sec, 52ºC for 30 sec, 72ºC for 1.5 min] and a final extension step of 5 min at 72ºC. PCR products were checked on 1% agarose gels and sequencing was performed at Macrogen, Inc. (Seoul, Korea) using the PCR primers. In all, we generated 44 sequences for 21 species. Several low copy genes (GBSS, GS2, and STG) did not amplify or showed multiple unspecific bands, indicating low primer specificity (Table 1). Others, such as GP1, were successfully amplified in a pilot study of a few individuals, but later sequencing indicated high heterozygosity that would require extensive subcloning. Only two regions, PHYC and EMB2765, were successfully amplified and sequenced in a majority of taxa. Hence, phylogenetic analysis and discussion of results were based on these regions.

2.3 Phylogenetic analysis

DNA sequences were edited using Sequencher 4.7 (Gene Codes, Ann Arbor, MI). The alignment was done with the online version of MAFFT (Stamatakis et al. 2008, using the L-INS algorithm) and manually adjusted in the editor Se-Al v. 2.0a11 (Rambaut 2002); gaps were treated as missing data. The alignment is available from the corresponding author. Resulting matrices were analysed under Bayesian inference using MrBayes 3.2cvs (Ronquist et al. 2008), with two

parallel runs of four chains each for 2 million generations and sampling every 1000 generations. Nucleotide substitution models were chosen based on the Akaike Information Criterion (Akaike 1973) as implemented in MrModeltest 2.3 (Nylander 2004). The GTR model was selected for EMB2765, and the HKY model for PHYC, with rate variation among sites in both. We used the program Tracer v. 1.5 (Rambaut and Drummond 2003 – 2009) to verify that all the parameters had reached the stationary phase in log-likelihood values and the split frequency criterion in MrBayes to assess convergence among chains. The initial 35,000 generations were discarded as burn-in samples and the remaining trees pooled to estimate the posterior probability distribution of the phylogeny and Bayesian clade posterior probabilities (pp).

Evolutionary data is most often presented as a phylogenetic tree with the underlying assumption that evolution is a branching process. However, nuclear markers are subject to different types of recombination – in vivo by meiotic crossing over, i.e. inter-allele or inter-locus, and in vitro during PCR. The probability of observing recombination increases with gene copy number. Recombination results in conflicting phylogenetic signals within the sequence and cannot be expressed as a branching topology (Martin et al. 2011). We explored the use of network representations (split networks) to represent ambiguous signals in the dataset and to detect possible recombination events. We used the neighbour-net method implemented in the software SplitsTree 4.0 (Huson and Bryant 2006) and difference distance-based algorithms to analyse each marker separately. We also used the program RDP beta version 3.34 (Martin et al. 2010) to test for possible recombination events in the LCG markers. This software applies a number of recombination detection and analysis algorithms for detecting putative recombination breakpoints. We used all eight methods available, with a *p* value of 0.1 with Bonferroni correction to initially provide a low stringency examination of putative breakpoints. We did, however, require phylogenetic evidence for recombination, with internal reference sequences for RDP, with default options for the other methods.

3. Results and discussion

3.1 Phylogenetic utility of DNA sequence loci

Not every taxon analysed could be sequenced for all markers. Most technical difficulties were apparently related to low primer specificity owing to base mismatches in the primer site. The amplification sometimes yielded multiple bands and after sequencing we occasionally found polymorphic sites. We excluded specimens with multiple signals excepting those with single polymorphic sites; in such cases we created two sequences. Split networks showed a small number of contradictory characters (represented by a box in the figure) in all three markers analysed (Fig. 1–3a), which could be an indication of homoplasy or recombination. RDP2, however, did not detect significant evidence of recombination in any of the markers: the two events detected in PHYC by a single method each were only at $p > 0.05$, and therefore disregarded.

Levels of sequence variation varied between loci (Table 3). ITS had the highest number of parsimony informative characters and number of variable characters. The nuclear exon EMB2765 and the protein-coding PHYC exhibited similar levels of variation to one another, which exceed those reported in *Hypericum* for some commonly used fast-evolving chloroplast spacers (Meseguer et al. 2013). On the other hand, substitution rates for the LCGs were half those of ITS (Table 3). Meseguer et al. (2013) reported high rates of nucleotide substitution in the ITS ribosomal spacer, which made alignment, especially with outgroups, exceptionally difficult. Alignment was considerably more straightforward in EMB2765 and PHYC, which, together with good phylogenetic support and levels of resolution (see below), make these LCGs a good alternative to ITS in *Hypericum* phylogenetic inference. This agrees with Wurdack and Davis (2009), who found PHYC and EMB2765 to be useful markers for sequencing across a range of Malphigiales and more distant outgroups, although the authors excluded the third codon position in their analyses (excluding the third codon position did not change the results of our analyses).

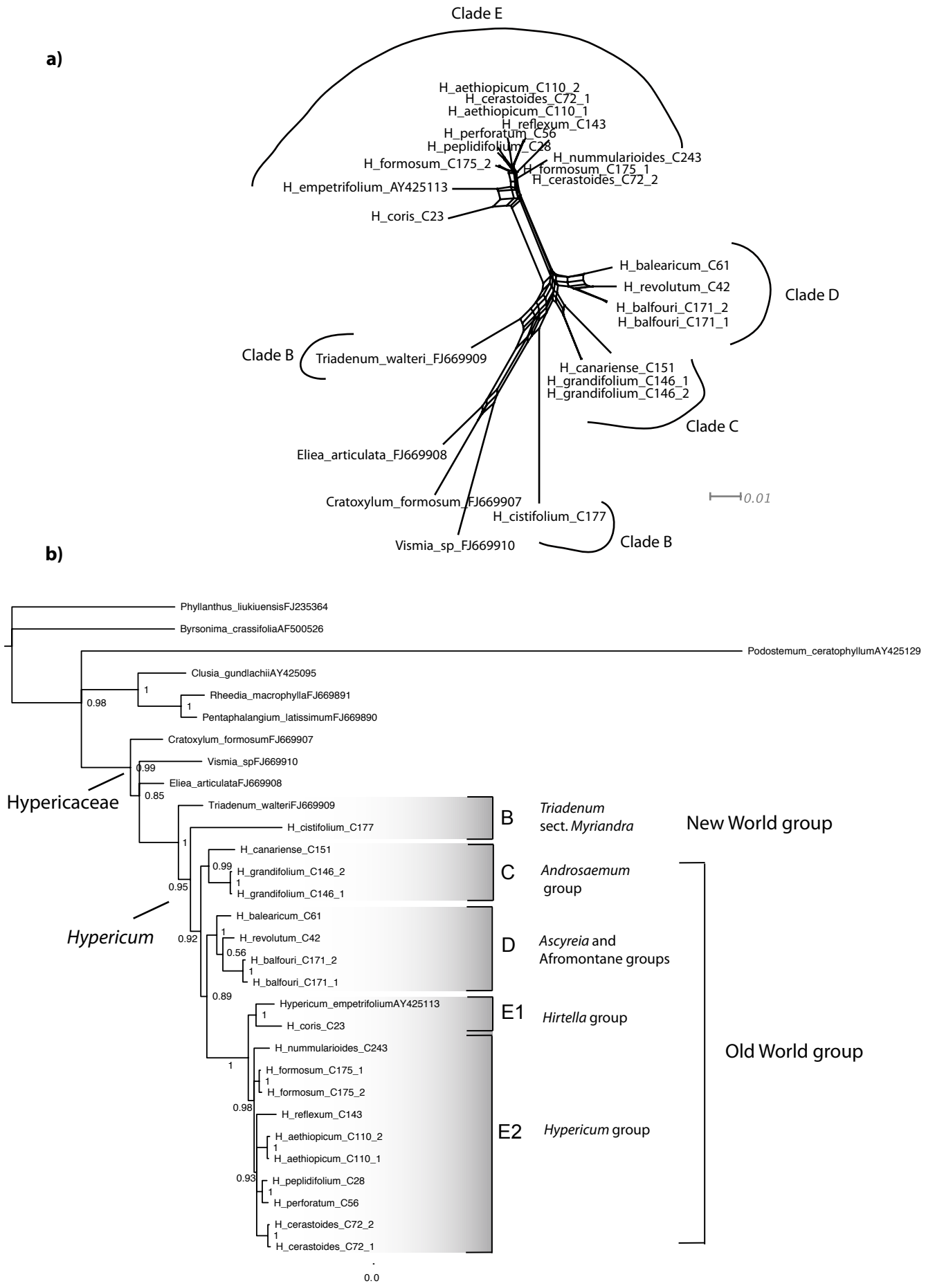


Figure 1. a) PHYC phylogenetic network in the genus Hypericum. Weight threshold=0.001. b) Phylogenetic relationships in Hypericum and related taxa inferred from the nuclear PHYC marker. 50% Bayesian Majority-Rule consensus tree showing posterior probabilities. A to E letters indicate major clades as defined in Meseguer et al. (2013).

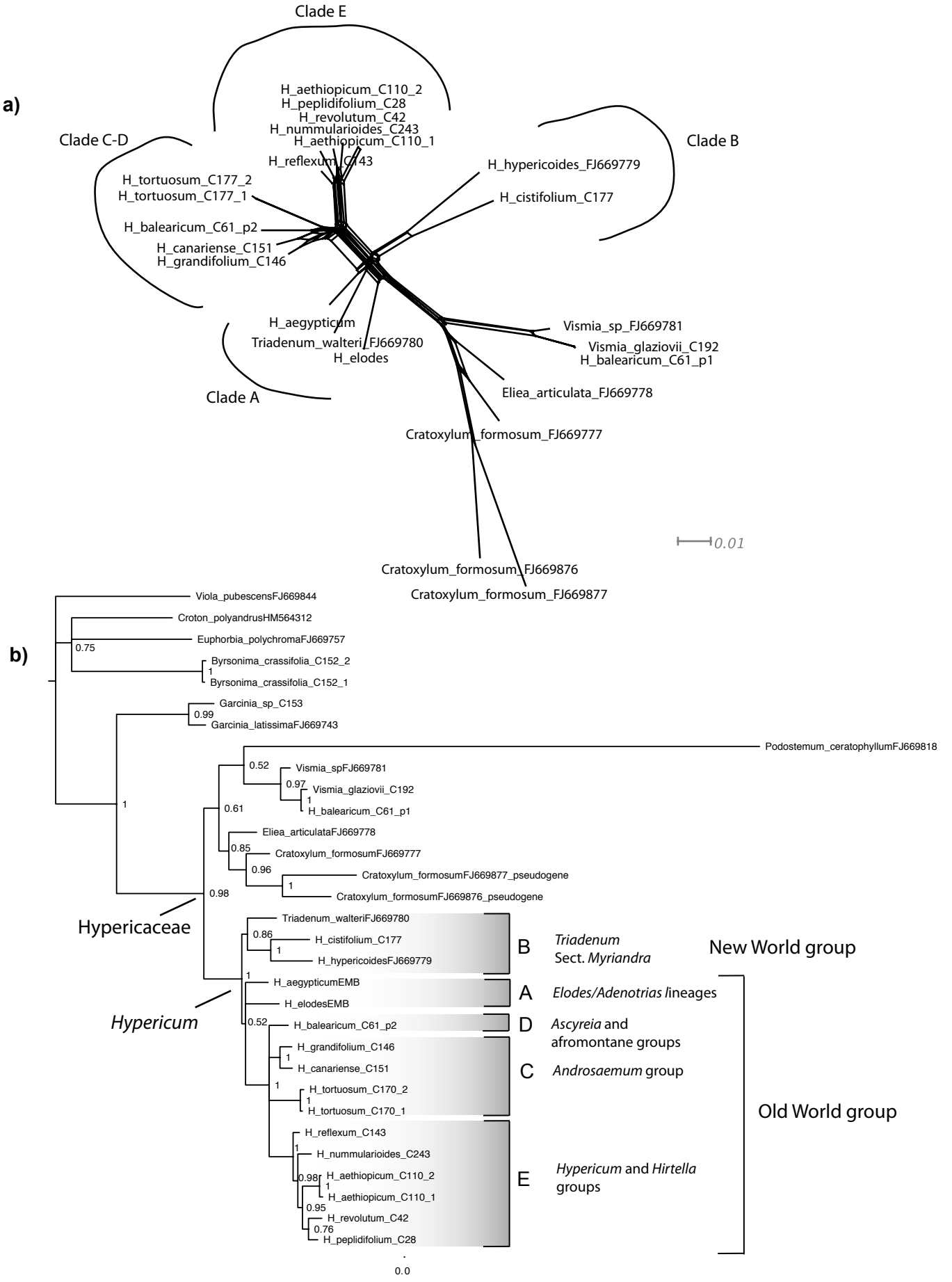


Figure 2. a) EMB2765 phylogenetic network in the genus Hypericum. Weight threshold=0.001. b) Phylogenetic relationships in Hypericum and related taxa inferred from the nuclear EMB2765 marker. 50% Bayesian Majority-Rule consensus tree showing posterior probabilities. A to E letters indicate major clades as defined in Meseguer et al. (2013).

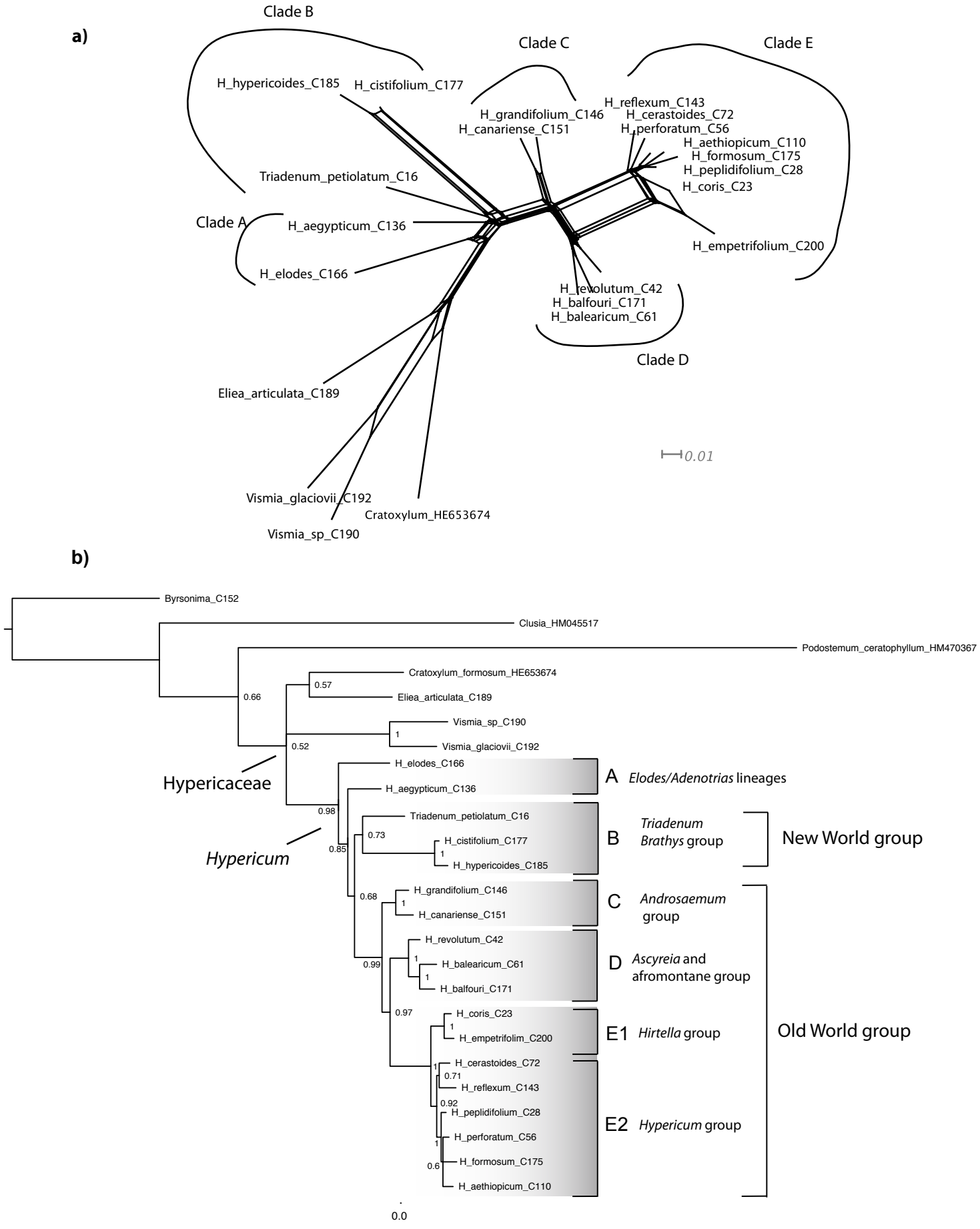


Figure 3. a) ITS phylogenetic network in Hypericum. Weight threshold=0.001. b) Phylogenetic relationships in the genus Hypericum and related taxa inferred from the nuclear ITS marker. 50% Bayesian Majority-Rule consensus tree showing posterior probabilities. A to E letters indicate major clades as defined in Meseguer et al. (2013).

Character consistency within each marker was calculated using the consistency, homoplasy and retention indices (CI, HI and RI, respectively) calculated on the Bayesian topologies (Table 3). This showed that the RI within EMB2765 is comparable to that found in ITS, whereas the CI and HI are more favourable in EMB2765 than in ITS. Within PHYC, all three indices have scores that are considerably better than those found in ITS (Table 3).

3.2 Phylogenetic relationships

The Bayesian 50% majority rule consensus trees for PHYC, EMB2765 and ITS are given in Figures 1b, 2b and 3b, respectively. Comparison among the three phylogenies is not straightforward because of the failure to amplify some taxa for several markers. However, the LCG phylogenies (Figs. 1b-2b) show a general topology that is largely congruent with the ITS tree (Fig. 3b) and with those reported in previous molecular studies of *Hypericum* (Rufhel et al. 2011; Nürk et al. 2013; Meseguer et al. 2013: Park and Kim 2004; Crocket et al. 2004). Network analysis showed similar groupings (Figs. 1a-3a). Among the LCG, resolution and clade support values were highest in the PHYC phylogeny (Fig. 1), which even showed higher support values for some basal nodes than those of other ITS phylogenies (Meseguer et al. 2013, Nürk et al. 2013; Park and Kim 2004; Crocket et al. 2004), or reported from concatenate plastid phylogenies (Meseguer et al. 2013). The PHYC phylogeny (Fig. 1) shows a geographic dichotomy between the "New World" lineage (*H. cistifolium*) and an Old World lineage, the latter divided into several clades that generally correspond to those found by Meseguer et al. (2013) and Nürk et al. (2013): clade C (*Androsaemum* group), clade D (*Ascyreia* and Afromontane groups) and clade E. Two subclades can be distinguished within clade E, the *Hirtella* group (clade E1) and the *Hypericum* group (clade E2, Meseguer et al. 2013), although interspecies relationships within the latter were not resolved. The EMB phylogeny showed generally lower support values for basal relationships but slightly better resolution at species-level (clade E2). This marker also recovers *Triadenum* as part of the New World lineage, confirming other phylogenetic studies that found *Hypericum* to be non-monophyletic (Meseguer et al. 2013, Nürk et al. 2013; Ruhfel et al. 2011). In contrast, the PHYC phylogeny shows *Triadenum* as the sister group of *Hypericum*, although this relationship is not well supported. Other incongruences between the two LCG markers affect phylogenetic relationships at the tribal level. In PHYC, the family Hypericaceae and the tribe Hypericeae are both recovered as well-supported monophyletic groups, whereas tribe Cratoxyleae (*Cratoxylum formosum* and *Eliea articulata*) appears as non-monophyletic (Fig. 1). In the EMB2765 tree, tribe relationships are better resolved, with tribe Hypericeae sister to Cratoxyleae-Vismieae. However, one species of *Hypericum*, *H. balearicum* appears within tribe Vismieae, as sister-group to *Vismia* (Fig. 2). The anomalous position of this species in the EMB2765 tree could be attributed to incomplete lineage sorting or to an ancient duplication event in which a different paralogue has been amplified in *H. balearicum*. The long branch separating the specimen *H. balearicum*_C61_p1 from the rest of *Hypericum*, and the short branch between this species and the outgroup *Vismia*, supports the later explanation as the most probable. To solve this issue, we designed a new internal primer for EMB based on the right *Hypericum* sequences to recover the orthologous copy of this marker in *H. balearicum*. The new sequence *H. balearicum*_C61_p2 fell into a congruent position with the other markers in the phylogeny (Fig. 2), lending support to the hypothesis that EMB2765 is present in more than one copy in *Hypericum*. Only diploid specimens have been described in *H. balearicum* (Robson, 1985), suggesting that heterozygosity probably owed to local duplications (paralogues). Further cloning strategies could help to do a better screening of gene copies. Nevertheless, our study suggests that the design of copy-specific primers can also be a useful strategy to address paralogy. Interestingly, Wurdack and Davis (2002) found non-functional pseudogene copies of EMB in *Cratoxylum* (Hypericaceae, Fig. 2), but these shorter copies grouped with the functional, full-length sequence (and were isolated using the same primers), suggesting a more recent duplication event.

Another incongruence between the phylogenies of EMB2765 and the other markers concerns the sister-group relationship of genus *Podostemum* with Vismieae, which renders family Hypericaceae not monophyletic. The long

branch subtending *Podostemum* suggests a potential "long branch attraction" artefact (LBA). This occurs when rapidly evolving lineages are inferred to be closely related, regardless of their true relationships (Bergsten 2005), and it is often observed among outgroup taxa misplacing long branched ingroup taxa. Although statistical, model-based approaches like Bayesian inference correct for multiple substitutions at the same site and are therefore more robust to LBA than parsimony, these methods are still susceptible to LBA-artefacts (Bergsten 2005). In this case, LBA is probably caused by a fast evolutionary rate in Podostemaceae, probably related to its switch to aquatic mode of life and extreme morphological modifications (Kita and Kato 2001), so adding more outgroup taxa to break up the long branch (Aguinaldo et al. 1997) could help here to solve the issue.

4. Conclussions

Our study shows the potential of two LCGs, EMB2765 and PHYC, for reconstructing phylogenetic relationships in genus *Hypericum* and related clades. A lower mutation rate in these markers in comparison with the ITS ribosomal spacer makes it easy to establish homologies in the alignment with outgroups, while pilot phylogenetic studies showed improved resolution and clade support values for basal nodes compared with ITS and other fast-evolving plastid markers. Further, the internal character consistency of the new markers is comparable to, or better than, that found in ITS. In this study, we have also discovered a paralogous copy in EMB2765 for *Hypericum* that was isolated through the design of copy-specific primers. This finding also constitutes a starting point to understand the complex dynamic of gene families.

References

Aguilar JF, Rosselló JA, Feliner GN (1999) Molecular evidence for the compilospecies model of reticulate evolution in Armeria (Plumbaginaceae). Systematic Biology 48:735–754.

Aguinaldo AMA, Turbeville JM, Linford LS, Rivera MC, Garey JR, Raff RA, Lake JA (1997) Evidence for a clade of nematodes, arthropods and other moulting animals. Nature, 387:489–493.

Akaike H (1973) Information theory and an extension of the maximum likelihood principle. In: Kiado A (ed). Second International Symposium on Information Theory. Budapest, pp 267–281.

Ané C, Larget B, Baum DA, Smith SD, Rokas A (2007) Bayesian estimation of concordance among gene trees. Molecular Biology and Evolution 24:412–426.

Bais HP, Vepachedu R, Lawrence CB, Stermitz FR, Vivanco JM (2003) Molecular and biochemical characterization of an enzyme responsible for the formation of hypericin in St. John's wort (*Hypericum perforatum* L.). J Biol Chem 278:32413–32422.

Ballard JW (2000) When one is not enough: introgression of mitochondrial DNA in Drosophila. Mol Biol Evol 17:1126–1130.

Bergsten J (2005) A review of long-branch attraction. Cladistics 21:163-193.

Blanco-Pastor JL, Vargas P, Pfeil BE (2012) Coalescent Simulations Reveal Hybridization and Incomplete Lineage Sorting in Mediterranean *Linaria*. PlosOne 7:e39089

Bloomquist EW, Suchard MA (2010) Unifying Vertical and Nonvertical Evolution: A Stochastic ARG-based Framework. Systematic Biology 59:27–41.

Crawford DJ, Mort ME (2004) Single-locus molecular markers for inferring relationships at lower taxonomic levels: observations and comments. Taxon 53:631–635

Clegg MT, Cummings MP, Durbin ML (1997) The evolution of plant nuclear genes. Proc Natl Acad Sci USA 92:7791–7798.

Crockett SL, Douglas AW, Scheffler BE, Khan IA (2004) Genetic profiling of *Hypericum* (St. John's wort) species by nuclear ribosomal ITS sequence analysis. Pl Med 70:1–7.

Denton AL, McConaughy BL, Hall BD (1998) Usefulness of RNA polymerase II coding sequences for estimation of green plant phylogeny. Mol Biol Evol 15:1082–5.

Doyle JJ (1992) Gene trees and species trees: molecular systematics as one-character taxonomy. Systematic Botany 17:144–163.

Ekenäs C, Heidari N, Andreasen K (2012) *Arnica* (Asteraceae) phylogeny revisited using RPB2: Complex patterns and multiple d-paralogues. Mol Phylogenet Evol 64:261–70

Frajman B, Eggens F, Oxelman B (2009) Hybrid origins and homoploid reticulate evolution within Heliosperma (Sileneae, Caryophyllaceae)-A multigene phylogenetic approach with relative dating. Systematic Biology 58:328–345.

Gatesy J, O'Grady P, Baker R (1999) Corroboration among data sets in simultaneous analysis: Hidden support for phylogenetic relationships among higher level artiodactyl taxa. Cladistics 15:271-313.

Heled J, Drummond AJ (2010) Bayesian Inference of Species Trees from Multilocus Data. Molecular Biology and Evolution 27:570–580.

Holland B, Benthin S, Lockhart P, Moulton V, Huber K (2008) Using supernetworks to distinguish hybridization from lineage-sorting. BMC Evolutionary Biology 8:202.

Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

Innes RW, Ameline-Torregrosa C, Ashfield T, Cannon E, Cannon SB, Chacko B, Chen NWG, Couloux A, Dalwani A, Denny R, et al (2008) Differential accumulation of retroelements and diversification of NB-LRR disease resistance genes in duplicated regions following polyploidy in the ancestor of soybean. Plant Physiology 148:1740–1759

Joly S, McLenachan PA, Lockhart PJ (2009) A statistical approach for distinguishing hybridization and incomplete lineage sorting. American Naturalist 174:E54–E70

Jones G, Sagitov S, Oxelman B (2013) Statistical inference of allopolyploid species networks in the presence of incomplete lineage sorting. Systematic Biology 62:1–12.

Karppinen K, Hohtola A (2008) Molecular cloning and tissue-specific expression of two cDNAs encoding polyketide synthases from Hypericum perforatum. J Plant Physiol 165:1079–1086.

Kay KM, Whittall JB, Hodges SA (2006) A survey of nuclear ribosomal internal transcribed spacer substitution rates across angiosperms: an approximate molecular clock with life history effects. BMC Evolutionary Biology 6:36 doi:10.1186/1471-2148-6-36.

Kita Y, Kato M (2001) Infrafamilial Phylogeny of the Aquatic Angiosperm Podostemaceae Inferred from the Nucleotide Sequences of the matK Gene. Plant Biology 3:156–163.

Kubatko LS, Degnan JH (2007) Inconsistency of phylogenetic estimates from concatenated data under coalescence. Systematic Biology 56:17–24.

Legendre P, Makarenkov V (2002) Reconstruction of biogeographic and evolutionary networks using reticulograms. Systematic Biology 51:199–216.

Liu B, Falkenstein-Paul H, Schmidt W, Beerhues L (2003) Benzophenone synthase and chalcone synthase from *Hypericum androsaemum* cell cultures: cDNA cloning, functional expression, and site-directed mutagenesis of two polyketide synthases. Plant J 34:847–855.

Mallet J (2007) Hybrid speciation. Nature 446:279–283.

Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefeuvre P (2010) RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26:2462-2463.

Martin DP, Lemey P, Posada D (2011) Analysing recombination in nucleotide sequences. Molecular Ecology Resources 11:943–955.

Mayol M, Rossellò JA (2001) Why nuclear ribosomal DNA spacers (ITS) tell different stories in Quercus?. Mol Phylogenet Evol 19:167–176.

Maureira-Butler IJ, Pfeil BE, Muangprom A, Osborn TC,

Doyle JJ (2008) The reticulate history of Medicago (Fabaceae). Systematic Biology 57:466–482.

Meng C, Kubatko LS (2009) Detecting hybrid speciation in the presence of incomplete lineage sorting using gene tree incongruence: A model. Theoretical Population Biology 75: 35-45

Meseguer AS, Aldasoro JJ, Sanmartín I (2013) Bayesian inference of phylogeny, morphology and range evolution reveals a complex evolutionary history in St John's wort (*Hypericum*). Molecular Phylogenetics and Evolution 67:379-403.

Nieto-Feliner G, Rosselló JA (2007) Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. Molecular Phylogenetics and Evolution 44:911–919.

Nürk NM, Blattner FR (2010) Cladistic analysis of morphological characters in *Hypericum* (Hypericaceae). Taxon 59:1495-1507.

Nürk NM, Madriñán S, Carine MA, Chase MW, Blattner FR (2013) Molecular phylogenetics and morphological evolution of St. John's wort (*Hypericum*; Hypericaceae). Molecular Phylogenetics and Evolution 66:1-16.

Nylander JAA (2004) MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Oxelman B, Yoshikawa N, McConaughy BL, Luo J, Denton AL, Hall BD (2004) RPB2 gene phylogeny in flowering plants, with particular emphasis on asterids. Mol Phylogenet Evol 32:462–79.

Park SJ, Kim KJ **(**2004) Molecular phylogeny of the genus *Hypericum* (Hypericaceae) from Korea and Japan: Evidence from nuclear rDNA ITS sequence data. Journal of Plant Biology 47:366–374.

Pfeil BE (2009) The effect of incongruence on molecular dates. Taxon 58:511-518

Pfeil BE, Brubaker CL, Craven LA, Crisp MD (2004) Paralogy and orthology in the MALVACEAE rpb2 gene family: investigation of gene duplication in hibiscus. Mol Biol Evol 21:1428-37.

Pilepić KH, Balić M, Blažina N (2011) Estimation of phylogenetic relationships amongsome *Hypericum* (Hypericaceae) species using internal transcribed spacer sequences. Plant Biosystems 145:81–87.

Rambaut A, Drummond AJ (2003 – 2009) Tracer, version 1.5, MCMC trace analysis package. Website http://tree.bio.ed.ac.uk/software/.

Rambaut A (2002) Se-Al: Sequence Alignment Editor. Available from: http://tree.bio.ed.ac.uk/software/seal/.

Rauscher JT, Doyle JJ, Brown AH (2002) Internal transcribed spacer repeat-specific primers and the analysis of hybridization in the Glycine tomentella (Leguminosae) polyploid complex. Mol Ecol 11:2691-702.

Robson NKB (1977) Studies in the genus *Hypericum* L. (Guttiferae) - 1. Infrageneric classification. Bull Brit Mus (Nat Hist), Bot 5:295–355.

Robson NKB (1981) Studies in the genus *Hypericum* L. (Guttiferae). 2. Characters of the genus. Bull Brit Mus (Nat Hist), Bot 8:55–226.

Robson NKB (1985) Studies in the genus *Hypericum* L. (Guttiferae). 3. Sections 1. *Campylosporus* to 6a. *Umbraculoides*. Bull Brit Mus (Nat Hist), Bot 12:163–211.

Robson NKB (2012) Studies in the genus *Hypericum* L. (Hypericaceae) 9. Addenda, corrigenda, keys, lists and general discussion. Phytotaxa 72:1–111.

Ruhfel BR, Bittrich V, Bove CP, Gustafsson MHG, Philbrick CT, Rutishauser R, Xi Z, Davis CC (2011) Phylogeny of the clusioid clade (Malpighiales): Evidence from the plastid and mitochondrial genomes. American Journal of Botany 98:306–325.

Sang T (2002) Utility of Low-Copy Nuclear Gene Sequences in Plant Phylogenetics. Critical Reviews in Biochemistry and Molecular Biology 37:121–147

Small RL, Cronn RC, Wendel JF (2004) L. A. S. Johnson review no. 2. Use of nuclear genes for phylogeny reconstruction in plants. Austral Syst Bot 17:145–170.

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web servers. Systematic Biology 57:758–771.

Stefanovic S, Pfeil BE, Palmer JD, Doyle JJ (2009) Relationships among phaseoloid legumes based on sequences from eight chloroplast regions. Syst Bot 34:115–128.

Stevens PF (2007) Hypericaceae. In: Kubitzki K (ed). The families and genera of vascular plants. Springer, Berlin, pp 194–201.

Wendel JF, Schnabel A, Seelanan T (1995) Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). Proc Natl Acad Sci USA 92:280–284.

White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR protocols: a guide to methods and applications, 315–322.

Wurdack KJ, Davis CC (2009) Malpighiales phylogenetics: Gaining ground on one of the most recalcitrant clades in the angiosperm tree of life. American Journal of Botany 96:1551–1570.

Yu Y, Cuong T, Degnan JH, Nakhleh L (2011) Coalescent Histories on Phylogenetic Networks and Detection of Hybridization Despite Incomplete Lineage Sorting. Systematic Biology 60:138–149.

TABLES

Table 1. Species included in this study and GenBank accession numbers. Morphological classification of *Hypericum* is based on Robson (1977–2010). The symbol * denotes sequences obtained from GenBank. Herbaria acronyms follow the abbreviations published in the Index Herbariorum.

Species	ID	Morphological section	Voucher	Genebank accessions		
				PHYC	EMB	ITS
Hypericeae						
H. aegypticum	C136	*Adenotrias*	7706 (GB)	XX00000	XX00000	KC709380
H. aethiopicum	C110	*Adenosepalum*	Aedo 14946 (MA)	XX00000	XX00000	KC709367
H. balearicum	C61	*Psorophytum*	Sanchez 13 (MA)	XX00000	XX00000	KC709338
H. balfouri	C171	*Campylosporus*	Aldasoro 14697 (MA)	XX00000	XX00000	KC709397
H. canariense	C151	*Webbia*	Aldasoro 10312 (MA)	XX00000	XX00000	KC709389
H. cerastoides	C72	*Campylopus*	727854 (MA)	XX00000	XX00000	KC709341
H. cistifolium	C177	*Myriandra*	Miller 8393 (MO)	XX00000	XX00000	KC709402
H. coris	C23	*Coridium*	Sanchez 5.1 (MA)	XX00000	XX00000	KC709429
H. elodes	C166	*Elodes*	Devain s.n. (MA)	XX00000	XX00000	KC709393
H. empetrifolium	C200	*Coridium*	Ruiz s.n. (MA)	-	-	KC709416
H. empetrifolium	*	*Coridium*	Chase 837 (K)	AY425113	-	-
H. formosum	C175	*Hypericum*	Merrill 12606 (MO)	XX00000	XX00000	KC709400
H. grandiflorum	C146	*Adenosepalum*	Aldasoro 10354 (MA)	XX00000	XX00000	KC709385
H. hypericoides	C185	*Myriandra*	Miller 8447 (MO)	-	-	KC709407
H. hypericoides	*	*Myriandra*	Wurdack D492 (US)	-	FJ669779	-
H. nummularioides	C243	*Taeniocarpum*	Sanchez 164 (MA)	XX00000	XX00000	-
H. peplidifolium	C28	*Humisfusoideum*	Aldasoro 10431 (MA)	XX00000	XX00000	XX00000
H. perforatum	C56	*Hypericum*	Tauleigne s.n. (MA)	XX00000	XX00000	KC709333
H. reflexum	C143	*Adenosepalum*	Aldasoro 10352 (MA)	XX00000	XX00000	KC709382
H. revolutum	C42	*Campylosporus*	Castroviejo 17247 (MA)	XX00000	XX00000	XX00000
H. tortuosum	C170	*Triadenioides*	Aldasoro 14645 (MA)	-	XX00000	-
Triadenum walterii	*	-	Brant 4792 (MO)	FJ669909	FJ669780	-
Triadenum petiolatum	C16	-	Correll 35026 (S)	-	-	KC709312
Vismieae						
Vismia glaziovii	C192	-	Niangadouma 374 (MO)	XX00000	XX00000	KC709411
Vismia	C190	-	Fuentes 10934 (MO)	XX00000	XX00000	KC709410
Vismia sp.	*	-	Miller 9313 (MO)	FJ669910	FJ669781	-
Cratoxyleae						
Cratoxylum formosum	*	-	Chase 1218 (K)	FJ669907	FJ669777	-
Cratoxylum formosum	*	-	Larson 33255 (B)	-	-	HE653674
Eliea articulata	*	-	Razakamalala 295 (MO)	FJ669908	FJ669778	KC709409
Outgroups						
Byrsonima crassifolia	*	-	Davis et al. (2002)	AF500526	-	-
Byrsonima sp.	C152	-	Aldasoro 9931 (MA)	XX00000	XX00000	XX00000
Clusia gundlachii	*	-	Chase 341 (NCU)	AY425095	-	-
Croton_polyandrus	*	-	Nunes 1376 (HUEFS)	-	HM564312	-
Euphorbia_polychroma	*	-	Chase 102 (NCU)	-	FJ669757	-
Garcinia sp.	C153	-	Aldasoro 9930 (MA)	XX00000	XX00000	-
Garcinia_latissima	*	-	Chase 2100 (K)	-	FJ669743	-
Pentaphalangium	*	-	Chase 1219 (K)	FJ669891	-	-
Phyllanthus liukiuensis	*	-	Kawakita 49	FJ235364	-	-
Podostemum ceratophyllum	*	-	Philbrick 6285 (WCSU)	-	-	HM470367

Species	ID	Morphological section	Voucher	Genebank accessions		
				PHYC	EMB	ITS
Podostemum ceratophyllum	*	-	Cusick 30042 (NY)	AY425129	-	-
Rheedia macrophylla	*	-	Chase 341 (NCU)	AY425095	-	-
Viola_pubescens	*	-	Wells 4886 (US)	-	FJ669844	-
Clusia sp.	*	-	MG-2010	-	-	HM045517

Table 2. Low copy nuclear regions and sequences of primers screened in this study.

Region	Primer name	Sequence 5'-3'	Reference
Phytochrome C	PHYC_Hyp_1F	CCAGCCACCGACATACCTCAAG	Own
	PHYC_Hyp_1R	GTAAGCTCCGCCACTTGAC	Own
	PHYC-INT1F	CCAGCTACTGATATACCWCARGCTTC	Wurdack & Davis_2009
	PHYC-INTR	CCAGCTTCCATAAGGCTATCAGTRCT	Wurdack & Davis_2009
Chalcone synthetase	CHS_Hyp_1F	GGAAGAAGTCAGGAAGGCGCAG	Own
	CHS_Hyp_1R	GGTCTCAACGGTAAGCCCAG	Own
	CHS_Hyp_2F	ACCGTGATGGCCATCGGAAC	Own
	CHS_Hyp_2R	CCAAAAAGCACTCCCCACTCGA	Own
Chloroplast-expressed glutamine synthetase	GScp687f	GATGCTCACTACAAGGCTTG	Emshwiller & Doyle_1999
	GScp994r	AATGTGCTCTTTGTGGCGAAG	Emshwiller & Doyle_1999
	GScp853f	TTACYGAACAAGCTGGYGTTGT	Emshwiller & Doyle_1999
	GScp856r	AGSACAACRCCAGCTTGTTC	Emshwiller & Doyle_1999
GBSSI Granule-bound starch synthase	Wax1f	CTG GTG GAC TTG GTG ATG	Own
	Wax1r	GGC YCC CAT DTG RAA TCC TGT G	Own
	Wax2f	CCT GKC TGC TCT KGA RGC AC	Own
	Wax2f	CCT TGG CAA GWG GAG CRA TCT CS	Own
Beta-carotene hydroxylase mRNA	Chyb_1F	TTG GCA (GA)AT GGA GGG TGG AGA	Own
	Chyb_1R	GGC STA YAT GTT TGT MCA YGA YGG	Own
Embryo defective 2765	EMB2765ex9F2	TATCCAAATGAGCAGATTATGTGGGA	Wurdack & Davis_2009
	EMB2765ex9R	TTGGTCCAYTGTGCWGCAGAAGGRT	Wurdack & Davis_2009
	EMB_Hyp_3F	TGA TTC CAA AAT TGC CTT GAA G	Own
	EMB_Hyp_4F	TGT CCA AGG CRA TAG TTA CAG TTC TC	Own
	EMB_Hyp_3R	CCA GGA AGC TGT CCC ACA	Own
TAFII15, Salt tolerance during germination 1	STG_Hyp_1F	CATCCCTGTTGATGGGCTRT	Duarte et al., 2010
	STG_Hyp_1R	GAAATTTGTTGCAGADGTTGC	Duarte et al., 2010
	STG_Hyp_2F	CTTGGACAGATCATCCATNGTCA	Duarte et al., 2010
Glucose-6-phosphate isomerase (Marcussen primer)	GPI_Hyp_12_1F	CGTGGTGCCACTGTCTCT	Own
	GPI_Hyp_16_1R	AGTTGRTAAAAGCTRTGCTG	Own
	GPI_Hyp_12_2F	CAATATGGTTTTCCAGTTGTTGA	Own
	GPI_Hyp_16_2R	GTTCCAGGTTCACCGAARTC	Own
	GPI_Hyp_13_3F	AGGTGCTGCAAGCATTGAT	Own

Table 3. Character-status summary for the nuclear low copy markers PHYC and EMB2765, and for the nuclear intergenic spacer ITS. TL mean: mean of the total tree length estimated over the two independent Bayesian runs. CI = consistency index, HI = homoplasy index, RI = retention index.

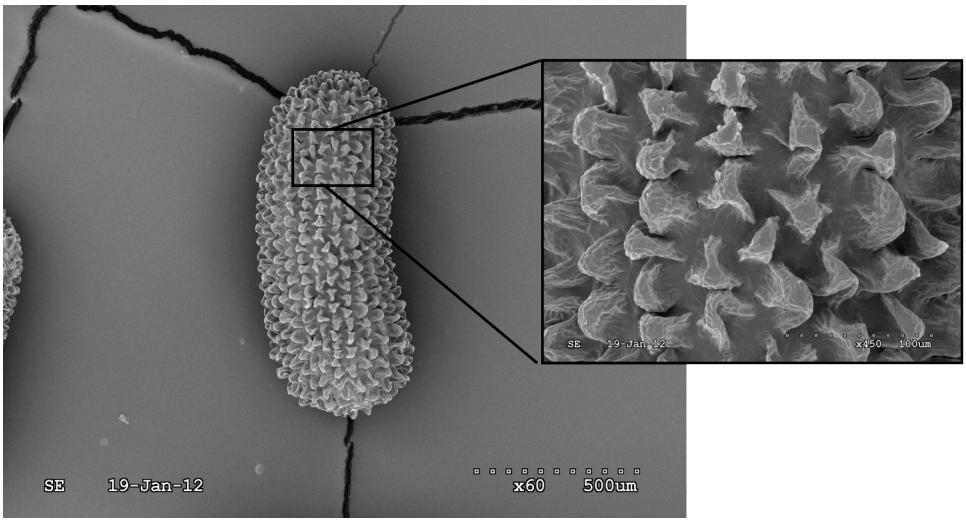
Region	# of characters	# parsimonious uninformative	# parsimonious informative	% informative characters	TL mean	CI	HI	RI
EMB2765	819	105	89	10.86	1.822462	0.782	0.218	0.749
PHYC	822	103	94	11.43	2.047945	0.853	0.147	0.879
ITS	740	105	164	22.16	3.929994	0.729	0.271	0.759

CHAPTER 3

Paleobiology of the genus *Hypericum* (Hypericaceae): a survey of the fossil record and its palaeogeographic implications

This chapter has been done in collaboration with Isabel Sanmartín from the Department of Biodiversity and Conservation, Real Jardín Bótanico-CSIC.

This chapter has been published in the peer-review journal *Anales del Jardín Botánico de Madrid*, volume 69, Issue 1, pages 97 to 106 (2012).



CHAPTER 3

Paleobiology of the genus *Hypericum* (Hypericaceae): a survey of the fossil record and its palaeogeographic implications

Andrea Sánchez Meseguer, Isabel Sanmartín

A b s t r a c t

Genus *Hypericum* is one of the 100 largest genera in angiosperms with nearly 500 species. Despite its worldwide, nearly cosmopolitan distribution and apparently old age – there are fossil remains of relatives from the Mid Cretaceous – the fossil record of *Hypericum* has been largely overlooked in phylogenetic studies. Here, we survey the fossil record of *Hypericum* from the literature, with special emphasis on the oldest fossil remain, *Hypericum antiquum*, from which we reassess its diagnostic characters. We evaluate the implications of this record in reconstructing the past geographic distribution of genus *Hypericum*.

1. Introduction

Hypericum L. is a large genus containing nearly 500 species (Nürk & Blattner, 2010; Robson, 2010). The genus comprises large shrubs, small trees or rhizomatous, some times annual, herbs, with yellow flowers and frequently glandulous tepals or leaves. It has a worldwide distribution, with representatives in nearly every continent, being only absent from the poles, deserts, and low-altitude tropical areas (Fig. 1). Nowadays, the largest diversity in the genus is found in the Northern Hemisphere (Eurasia and North America), but it is also abundant in high-altitude tropical areas of the Southern Hemisphere, such as the tropical An des in South America or the Eastern African Mountains (Robson, 1977). Some species of *Hypericum,* like *H. perforatum,* have been used in traditional medicine based on the pharmacological properties of their active compounds, hypericine and pseudohypericine, which are used as pain killers, antidepressants or anticancer treatments (Matzk & al., 2001).

The taxonomic adscription of the genus *Hypericum* has long been discussed, with the genus being classified as either a tribe or a subfamily inside a broadly defined family Clusiaceae (Robson, 1977). The most recent taxonomic treatment (Stevens, 2007) considers *Hypericum* as a genus inside the family Hypericaceae. This family includes three tribes: Hypericeae, which has a nearly cosmopolitan distribution and includes five genera (*Hypericum, Triadenum, Thornea, Santomasia,* and *Lianthus*), and the tropical tribes Cratoxyleae (*Eliea* and *Cratoxylon*) and Vismieae (*Harungana, Vismia* and *Psorospermum*). Recent phylogenetic revisions based on morphological (Nürk & Blattner, 2010) and molecular characters (Ruhfel & al., 2011) have shown that genus *Hypericum* is not monophyletic, and that other members of tribe Hypericeae (*Triadenum*, *Santomasia* and *Thornea*) are nested within it. Therefore, throughout this work we will refer to Hypericeae as a synonym to *Hypericum*. Although no attempt has been made yet to reconstruct the temporal evolution of *Hypericum* (Meseguer & al., in prep.), there is indirect evidence that the genus is rather old. The or der Malpighiales, to which family Hypericaceae belongs to – together with families Clusiaceae, Malpighiaceae and Bonnetiaceae, among others – began its diversification in the Early Cretaceous, as evidenced by the presence of a Mid-Cretaceous fossil, *Paleoclusia chevalieri* (Crepet & Nixon, 1998) and by molecular phylogenetic estimates of divergence times (Davis & al., 2005). The latter study estimates the age of the s plit of Hypericaceae and its sister group Podostemaceae around the Late Cretaceous, ca. 76 millions of years (Ma).

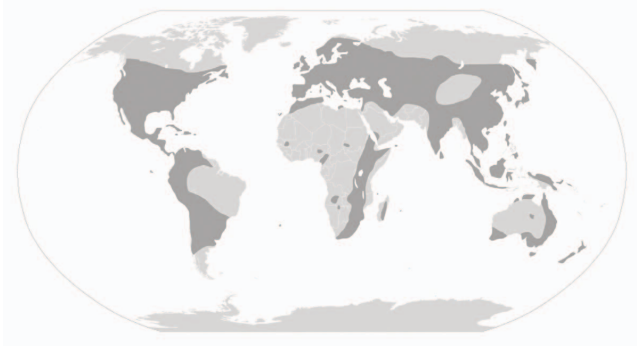


Fig. 1. Map showing the present distribution of *Hypericum* species (adapted from Robson, 1977).

Despite the presumably old age, worldwide distribution, and the ancient radiation of sister families, the fossil record of *Hypericum* has been mostly overlooked in phylogenetic studies. In what it is still the most exhaustive taxonomic revision of the genus, Robson (1981, pg. 65) wrote "In the absence of a useful fossil record – and with the possible exception of Pliocene seeds (see e.g. Reid, 1923), there are no known fossils of *Hypericum*…" Subsequent researchers have also assumed that *Hypericum* lacks a useful fossil record for phylogenetic purposes (Stevens, 2007). This is surprising, since a quick search in the Paleobiology Data base (http://paleodb.org) found nearly 70 collections (95 occurrences) from the Upper Eocene onwards. It should be noted, however, that many of these fossils are microfossils (seeds and pollen), and that the material is often fragmented or not well preserved, which makes difficult to as sign the fossil to a particular taxon.

Fossil remains can provide calibration points to estimate lineage divergence times from molecular-measured branch lengths in a phylogenetic analysis, the "so-called" molecular clock. In recent years, there have been important advances to incorporate the uncertainty in fossil calibrations ("stratigraphic ages") to phylogenetic dating (Ho, 2007). Likewise, fossils can be used to pinpoint the geographic location of now extinct ancestral lineages – which does not necessarily coincide with the distribution of the extant taxa – and this in turn can provide valuable information concerning the climatic preferences of the group. Here, we survey the fossil record of *Hypericum* from the literature, with special focus in the oldest fossil described (*H. antiquum* Balueva & V.P. Nikitin), and reassess some diagnostic char acters used for the taxonomic assignment to extant taxa, as a first step to employ these fossils in reconstructing the spatiotemporal evolution of the genus (Meseguer & al., in prep.).

2. Material and Methods

We searched the available paleobiological literature for *Hypericum* and other genera of tribe Hypericeae, especially through the online resource Paleobiology Database (http://paleodb.org), as well as other original publications describing fossil taxa of *Hypericum* not included in the re ferred database. In addition, to assess the main diagnostic features of seeds of extant *Hypericum* species, we used a Zeiss stereomicroscopy attached to a digital camera ("STE MI 2000CZEISS) for describing and documenting fine details of specimens preserved in MA. Microscopic features of the seed surface were photographed with a Hitachi S 3000N digital Scanning Electron Microscope at the facilities of the Real Jardín Botánico de Madrid, CSIC; prior to visualization, seed specimens were treated using a Balzers SDC004 sputter coater for sample gold-coating.

3. Pollen and seed morphology

Before reviewing the fossil record of *Hypericum*, we will describe the morphology and main diagnostic features of the seeds and pollen of extant species of *Hypericum*, since these remains are the most common fossil records attributed to the genus.

3.1 Pollen morphology

Pollen grains in the genus *Hypericum* have been studied in several works (Khan, 1969; Thomas, 1970; Clarke, 1975; 1976, 1981; Barros & Ramos, 1984; Mártonfi & al., 2002). The standard type of the group is a regular 3zonocolporate with a microreticulate to reticulate ornamentation pattern. The only divergence within *Hypericum* from the standard tricolporate plan characteristic of angiosperms is a tendency of some species to produce grains with more than three apertures and loosed strict polarity (Clarke, 1981). Al though the presence of irregular pollen grains have been frequently cited within the genus (Clarke, 1975; Mártonfi & al., 2002), among regular grains a set of different basic types can be distinguished. Clarke (1981) described eleven distinct pollen types that are more or less uniform within the defined morphological sections. The only exceptions are sections *Ascyreia*, *Triadenioides*, *Adenosepalum*, *Hypericum* and *Hirtella* whose species present pollen from different morphotypes (Table 1).

Despite the discreteness of the different pollen types within *Hypericum*, the basic plan is not particularly distinctive from other angiosperm lineages, which makes fossil seeds a potentially more powerful source of paleobiological information (Clarke, 1981). On the other hand, fossil pollen is not as useful as seeds to associate ancestors to a particular place in space, since pollen grains can occasion ally travel long distances and be found far away from the species original distribution.

3.2. Seed morphology

In angiosperms, the seed coat is usually divided into a testa and a tegmen, and both can be further divided into different epidermal layers. The seeds of the family Hypericaceae are exotegmal, which means that the mechanical tissue is developed in the outer layer of the tegmen, and the cells develop into a palisade of tubular or radially elongated cells, with stellate-undulate or lobate facetes (Corner, 1976a, 1976b). According to Corner (1976a), the shape of the exotegmic cells in this family is very characteristic and found only in a few other families: Clusiaceae, Elatinaceae and Geraniaceae. Moreover, Hypericaceae (excluding *Psorospermum* with fleshy tegmen cells) and Clusiaceae seeds share a unique structure with large, lignified, and tabular, thick-walled stellate cells in the exotegmen. Nevertheless, Hypericaceae seeds are distinguishable from Clusiaceae seeds in that the former are generally smaller and do not present an arille.

The typical seeds in Hypericaceae (Fig. 2) are small, narrowly cylindric to ovoidcylindric or ellipsoid, with size within the genus *Hypericum* ranging from 1.5 to 0.3 mm (Robson, 1981). They also present brownish to blackish colour due to tannin contents. Within Hypericaceae seeds, the most similar to Hypericeae are those of tribe Vismieae. Yet, *Vismia* seeds are generally larger (>1 mm long) and present a different testa sculpturing (Mourão & Beltrati, 2001; Arteaga, 2007).

In some *Harungana* and *Vismia* species there are large, swollen, orange or black glands in the seed testa that are not present in the seeds of Hypericeae (Stevens, 2007). The seeds of the tribe Cratoxyleae described so far are cartilaginous winged with a peripheral vein, very different to those of *Hypericum* (Robson, 1981). In addition, the seeds of tribe Hypericeae have a particular wing venation, composition and disposition of appendages that are not found in any other member of Hypericaceae. The wing in this group is thin and papery and it is some times reduced to a carina, basal prolongations or an apiculus (Robson, 1981). The different *Hypericum* seeds have also a characteristic reticulate pattern in the testa sculpturing. The walls of the cells of the outer layer (exotesta) are more or less thickened (Ohlendorf, 1907). This pattern of wall thickening is probably the most distinctive character of *Hypericum* and can sometimes help to distinguish entire sections (Robson, 1981; see Table 2 and Fig. 2).

Table 1. Basic pollen types in *Hypericum* sections. Adapted from Clarke (1981).

Pollen type	Description	Morphological section
1	Grains: prolatespheroidal to subprolate. Endoaperture: a porus with very small lateral extensions. Ornamentation: a tectum perforatum or microreticulum; tectal perforations regularly spaced.	*Campylosporus, Webbia, Adenosepalum,, Ascyreia*
2	Grains: prolatespheroidal to subprolate. Endoaperture: a porus, often more or less lalongate (transversally elongated), with very small lateral extensions. Ornamentation: a tectum perforatum or microreticulum; tectal perforations grouped together.	*Ascyreia*
3	Grains: subprolate. Endoaperture: a lalongate colpus. Ornamentation: a tectum perforatum or microreticulum; tectal perforations grouped together.	*Ascyreia, Takasagoya, Roscyna, Inodorum, Androsaemum*
4	Grains: subprolate. Endoaperture: a lalongate colpus. Ornamentation: a tectum perforatum or microreticulum; tectal perforations regularly spaced.	*Androsaemum, Bupleuroides, Psorophytum, Arthrophyllum, Triadenioides, Origanifolia, Hypericum*
5	Grains: spheroidal or prolatespheroidal. Oultine in polar view triangular with concave side. Endoaperture: a lalongate colpus. Ornamentation: a tectum perforatum or microreticulum; tectal perforations regularly spaced	*Adenotrias*
6	Grains: perprolate or prolate. Endoaperture: a large porus, more or less circular or lolongate. Ornamentation: microreticulate or reticulate; lumina regularly spaced.	*Triadenioides*
7	Grains: very small, prolate. Endoaperture: a large lolongate (longitudinally elongated) porus. Ornamentation: microreticulate or reticulate; lumina regularly spaced.	*Myriandra*
8	Grains: prolate. Endoaperture: a very large lolongate porus or colpus. Ornamentation: microreticulate and reticulate; lumina regularly spaced.	*Trigynobrathys, Brathys*
9	Grains: very large, subprolate or prolate. Endoaperture: a more or less lalongate porus with short lateral and meridional extensions. Ornamentation: reticulate; lumina regularly spaced.	*Elodes*
10	Grains: subprolate or prolate. Endoaperture: a more or less lalongate porus with short lateral and meridional extensions. Ornamentation: microreticulate; lumina regularly spaced.	*Olympia, Campylopus, Drosocarpium, Oligostema, Thasia, Crossophyllum, Heterophylla, Hirtella, Coridium, Taeniocarpium, Hypericum, Concinna, Adenosepalum, Humifusoideum*
11	Grains all of irregular form and variable shape. Apertures varying in number from 2 to 12; arranged in many different ways.	*Hirtella*

In a recent cladistic analysis of *Hypericum*, Nürk & Blattner (2010) used some of the seed morphological characters discussed here to reconstruct phylogenetic relationships within *Hypericum* (Hypericeae), such as those related to the shape (cylindrical, fusiform, pear-shaped, ovoid, clavate, elongate), presence and type of appendages (carina, elaiosome, distal expansions, wings) and to the sculpturing of the testa (reticulate, scalariform, papillose).

In contrast to the exotesta, the endotesta in *Hypericum* is difficult to study and has rarely been described. Ramos Núñez (1983) studied the seeds of several Iberian species of *Hypericum*, and found that, while the exotesta allows distinguishing between sections, the endotesta sculpture can be specific for individual taxa. Most of the studied species exhibited a reticulate endotesta transversely oriented in relation with the longitudinal axis of the seed, which varied consistently in size, shape and surface of the cells (Ramos Núñez, 1983).

4. The fossil record of *Hypericum*

In this section, we reevaluate the taxonomic assignment of the oldest seed fossil remain attributed to *Hypericum, H. antiquum* from the Upper Eocene, using diagnostic characters of extant *Hypericum* seeds such as the major different sculpture patterns (Table 2).

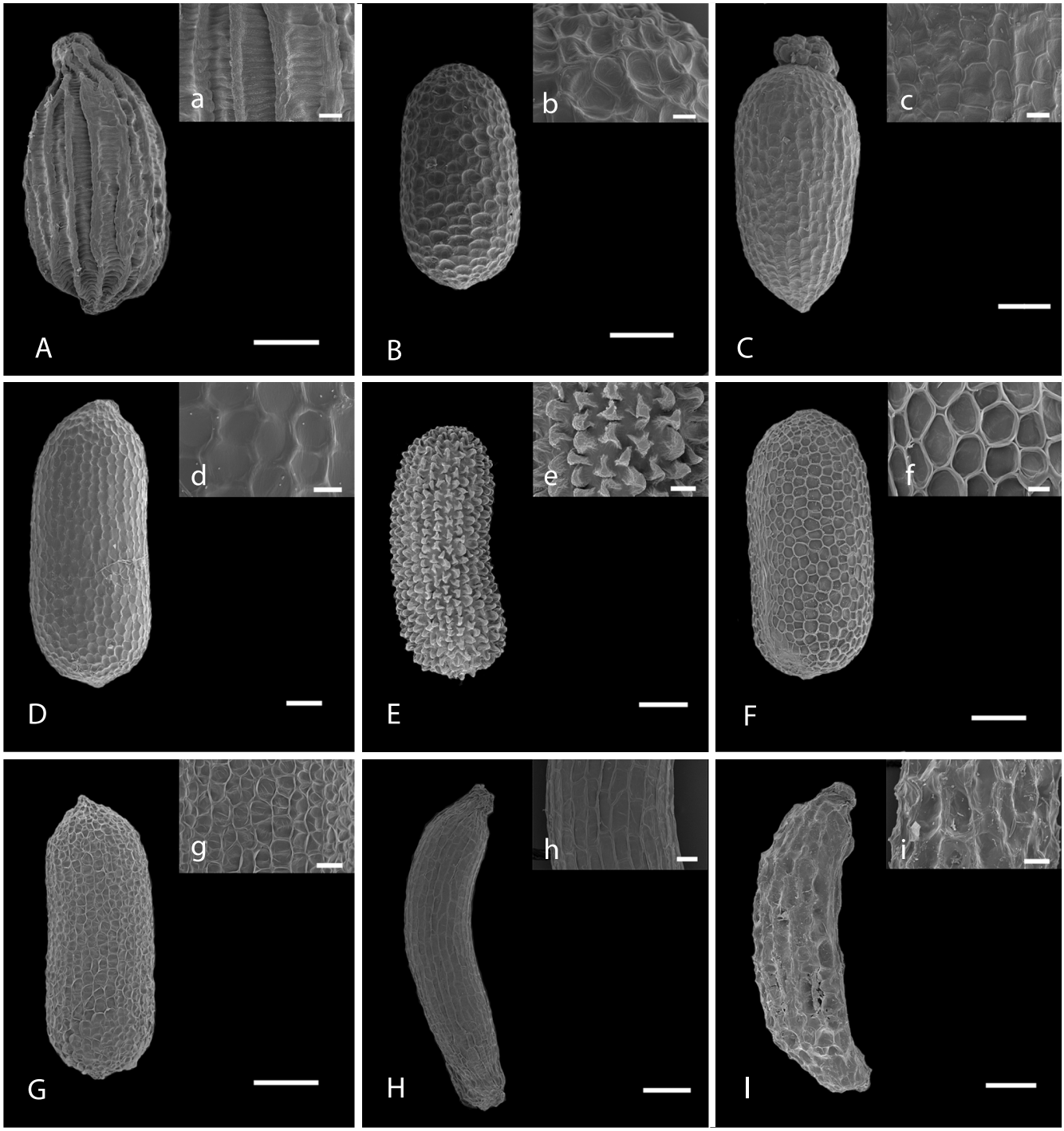


Fig. 2. Different *Hypericum* seeds with a detail of the sculpture: **A,** *H. elodes*; **B,** *H. peplidifolium*; **C,** *H. aegypticum*; **D,** *H. calycinum*; **E,** *H. hirsutum*; **F,** *H. perforatum*; **G,** *H. montanum*; **H,** *H. revolutum*; **I,** *H. canariensis* (AD; Fl = 200 µm; E= 400 µm; a, ce, gi= 40 µm; b, f= 20 µm).

We also present a chronological list of fossil remains from different geological periods that have been assigned to genus *Hypericum*. Although this list does not intend to be complete, especially for the most recent periods, it presents a first overview of the paleobiological literature available for the genus. While some of these fossils have been assigned to extant taxa, it is important to keep in mind that many of these taxonomic identifications are based on a limited number of characters of a few organs (pollen, seeds), and that the material is often fragmented and not well preserved. Even for perfectly conserved material, convergence of morphological traits might potentially obscure the true phylogenetic relation ship between extinct and extant taxa. In an extensive study of the fossil floras of Central Europe, Mai (2001) reviewed several occurrences of *Hypericum* microfossils and provided a key to distinguish between seeds of some fossil (and extant) species. This key, however, was far from being exhaustive, and more importantly, it did not include some of the oldest fossil records assigned to this genus, such as *H. antiquum* or *H. septestum* (see below); thus, we do not include it here.

Hypericum antiquum Balueva & V.P. Nikitin

To our knowledge, the oldest described fossil remains that can be unequivocally assigned to genus *Hypericum* corresponds to the seeds of the extinct species *Hypericum*

an tiquum Balueva & Nikitin from the Upper Eocene (40.4-33.9 Ma) of West Siberia (Arbuzova, 2005).

Table 2. Patterns in the sculpture of the testa in the different *Hypericum* sections. Adapted from Robson (1981).

Testa pattern		Description	Morphological Sections
Reticulate	Linear-reticulate or reticulate (RE-Fig.2; H)	Exotestal cells form roughly defined lines and have relatively thin walls; often the lines are deformed making the testa merely reticulate.	*Campylosporus, Psorophytum, Ascyreia, Takasagoya, Androsaemum, Inodorum, Roscyna, Hypericum, Myriandra, Humifusoideum, Adenosepalum, Santomasia*
	Foveolate (FO-Fig.2; F) or linear foveolate	All the walls are thickened leaving a round depression between them	*Bupleuroides, Hypericum, Concinna, Olympia, Campylopus, Origanifolia, Drosocarpium, Oligostema, Thasia, Crossophyllum, Heterophyla, Webbia, Triadenioides, Ascyreia, Coridium, Taeniocarpium*
Scalariform	Scalariform-reticulate (S-RE-Fig.2; C)	Thickening confined to the longitudinal cell walls	*Trigynobratys, Myriandra, Humifusoideum, Adenosepalum, Hypericum*
	Ribbed-scalariform (RI-S-Fig.2; A)	Where this longitudinal thickening is more pronounce, the seed looks ridge. It is often accompanied or preceded by lateral elongation of the cells.	*Elodes, Brathys, Trigynobrathys, Drosocarpium*
Papillose	Rugulose (RU)	Where the outer cell walls of the foveolate testa have begun to protrude	*Arthrophyllum, Coridium, Taeniocarpium, Origanifolia, Adenotrias*
	Papillose (PA-Fig.2; E)	The outer cell walls have a more pronounced protrusion	*Hirtella*

Surprisingly, these fossils have gone unnoticed in the Western literature, probably because they were first described in Russian in a poorly distributed book. Since these fossils have important implications for the paleobiological history of the genus (see below), we provide here a full description translated directly from the original source and a discussion of the main diagnostic characters and affinities used to assign this species to other extant taxa (Fig. 3).

Holotype: seed, Upper Eocene, Uzhanikha, borehole 1, depth 250 m (West Siberia), specimen 18/1, collection Kpr.611250 [NG].

Locality. Uzhanikha, Novosibirskaya oblast, Russia.

Diagnosis. Seeds 0.40.65 × 0.250.35mm, anatropus, cylindrical, often flattened, longitudinally slightly bent to wards the raphe. Meshes of the surface narrow; longitudinally elongated, hexagonal. The elongated cells of the testa form a mesh sculpture of the surface. Their arrangement on the surface of the seed gives the impression of a cross hatched line. Longitudinal walls of the cells are higher than the transverse ones, forming thin meridional ribs on the seed surface. One end is rounded with a small tubercle slightly shifted to the ventral side. The other end is slightly narrowed with a little tubercle. The seed case is relatively thin. The seed is black.

The description above and the original drawing of the fossil seed seem to present a completely preserved exotesta (Fig. 3A), from which a more detailed view of the cells is also provided (Fig. 3B). No mention about the endotesta is given. The size, colour, shape and testa sculpturing of these fossils are all characteristic of the seeds of the genus *Hypericum* (Fig. 2). Yet, the author (Arbuzova, 2005) considers *H. antiquum* as a different species from other known Oligocene *Hypericum* fossils (e.g., *H. septestum*, see below) and did not assign it to any extant representative. However, we think that the characteristic sculpturing pattern of the testa, with meridional ribs on the seed surface (Fig. 3), resembles those of some extant species, especially the pattern RIS that is characteristic of the following sections: *Elodes*, *Brathys*, *Trigynobrathys* and *Drosocarpium* (Table 2). Nevertheless, we found some ambiguities in the original description: there are no meridional or longitudinal ribs in the de tail of the exotesta (Fig. 3B), and although the axis of elongation of the palisade cells is described as longitudinal in the text, it appears transversal in the original drawing (Fig.

3B). In a subsequent publication, Nikitin (2006) seemed to correct this error and described the seeds as “made up of transversely elongated cells”. We also could not find any in formation about the number of cells between ribs in the text.

4.1. Other fossils assigned to Hypericum

4.1.1. Paleogene

There is a rich record of different fossil taxa recorded for this period. An Early Paleocene (65.561.7 Ma) *Hypericum* seed fossil has been cited in NE China (Hao & al., 2010). However, this work did not include any morphological description or illustration of the fossil and it is difficult to con firm whether it belongs to *Hypericum* and not to any other related group. The seed remains known as *H. bornense* Mai in Mai & Walter (1978) and *H. septestum* Nikitin (Doro feev, 1957; Dorofeev, 1963; Mai & Walther, 1978; Mai, 1997) appear in different Early to

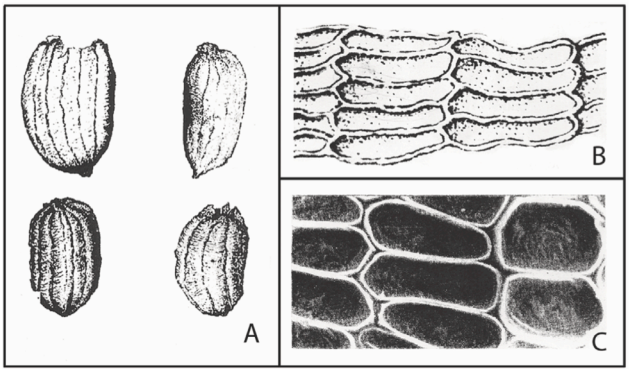


Fig. 3. A, fossil seeds of *Hypericum antiquum*; **B,** detail of the palisade cells of *H. antiquum* fossil seeds; **C,** palisade cells of the fossil specie *Triadenum virginicum* (= *H. virginicum*) (reproduced with permission from Arbuzova, 2005).

Late Oligocene sites (33.923 Ma) in Germany (Saxony) and in the Russian Federation (Tomsk). In particular, *H. septestum* seeds are a common element in the fossil record of *Hypericum* until the Pliocene, and have been described from different sites across Europe, including the Early Miocene (2316 Ma) of the Czech Republic (Teodoridis, 2002, 2003), Germany (Gümber & Mai, 2002), Poland (Holý, 1974; Kvaček & Teodoridis, 2003) and Russia (Dorofeev, 1963), the Mid Miocene (1611.6 Ma) of Bulgary (Palamarev & al., 2005) or the Pliocene (5.32.6 Ma) of Italy (Martinetto & al.,2007). The seeds of *H. septestum* are generally small, oval with a reticulate surface consisting of polygonal cells (Teodoridis, 2003). Remains attributed to this name have been variously assigned to different extant species. Negru (1972) and Dorofeev (1963) have compared them to *H. scabrum* L. and *H. aegypticum* L. respectively, whereas Gümbel & Mai (2002) and Arbuzova (2005) considered this species to be more similar to *Triadenum virginicum* L. (= *H. virginicum* L.).

Seeds of *Hypericum coriaceum* Nikitin have been de scribed from several sites from the Late Oligocene onwards in Russia and Belarus (Nikitin, 1948; Dorofeev, 1959, 1963, 1979; Yakubovskaya, 1982; Velichkevich, 1982, 1990). The identity of this paleotaxon has been put in doubt by Mai (2001), who suggests to exclude it from the *Hypericum* fossil record and assign it instead to genus *Ludwigia* (Onagraceae).

Fossil seeds of *Hypericum miocenicum* Dorof. emend. Mai are recorded from the Late Oligocene (28.423 Ma) of Saxonia and Lusatia (Germany), but also from the Late Miocene (11.65.3 Ma) of Poznan (Poland) and the Pliocene of Thuringia (Germany) (Mai, 2001). *Hypericum miocenicum* differs from the more frequent *H. septestum* in the shape of the surface cells being more quadrangular to rectangular with the wall of the foveoles thinner (Mai, 2001). Finally, fossil pollen of *Hypericum* (no other assignment) has been described form the Lower Oligocene (Sant Gallart Formation) in the Ebro Basin (Cavagnetto & Anadón, 1996).

4.1.2 Neogene

There are many fossil remains identified as *Hypericum* from this period (Miocene-Pliocene), with some of them assigned to present taxa. We believe, however, that the morphological affiliation of some of these records attributed to extant *Hypericum* species needs of further clarification. The seeds of *Hypericum holyi* Friis are cited in Central European sites from the Lower to Upper Miocene (235.3 Ma) (Friis, 1985; Mai, 1999; Mai, 2000; Meller & Hoffman, 2004). *Hypericum tertiaerum* Nikitin fossil seeds also appear during the same period in East and Central Europe and Siberia (Dorofeev, 1963; Łańcucka-Środoniowa & al., 1981; Van der Burgh, 1987; Baranowska-Zarzycka, 1988; Dyjor & al., 1992; Mai, 2000, 2001). This species is also frequently found in Pliocene sites from the same regions (Nikitin, 1957; Dorofeev, 1979; Jahn & al., 1984; Yaku bovskaya, 1984; Mai & Walther, 1988; Velichkevich, 1990; Mai, 2001; Gümbel & Mai, 2004; Arbuzova, 2005). Ac cording to Velichkevich & Zastawniak (2003) the EastAsiatic *H. japonicum* Blume and the North American *H. virginicum* L. and *H. tubulosum* Walt. are two extant species of *Hypericum* whose seeds most closely resemble those of *Hypericum tertiaerum*. Other Miocene *Hypericum* microfossils include: *H. tanaiticum* P. Dorof., *H. tambovicum* P. Dorof. in Russia, *H.* cf. *balearicum* L., *H.* cf. *scabrum* L. *H.* cf. *androsaemum* L. in Moldavia, and *H.* cf. *acutum* L. in Ukraine (Arbuzova, 2005); *H. welzowense* Mai in Germany (Mai, 2001), *H.* aff. *ponticum* Lipsky in Bulgaria (Palamarev & al., 2005), and *H.* cf. *humifusum* in northern Italy (KovarEder & al., 2006), as well as formally undescribed fossil remains from the Early Miocene of China (Zhao & al., 2004). In the Iberian Peninsula there are also pollen remains from the Mid Miocene (Barrón & al., 2010).

More recent records include the Pliocene seeds of *H. foveolatum* Dorof. from Russia (Dorofeev, 1986; Gümbel & Mai, 2004) and Belarus (Velichkevich & Zastawniak, 2003; Arbuzova, 2005), *H. androsaemun* L. *fossilis* (Gümbel & Mai, 2004), and fossil seeds of *H. perforatum* L. also from Russia (Arbuzova & al., 2005), which the authors mentioned to be identical to extant *H. perforatum* seeds. The fossil species *H. foveolatum* has been compared (Dorofeev, 1986) to extant species from Europe (*H. quadrangulum* L., *H. elegans* Willd., *H. tetrapterum* Fries), East Asia (*H. attenuatum* Choisy, *H. kamtschaticum* Ldb., *H. yezoënse* Maxim.), and North America (*H. nudiflorum* Michx., *H. microsepalum* Torr. & Grey), although none of these species have the testa surface cells morphologically identical to those of the fossil (Velichkevich & Zastawniak, 2003). Other fossils include a macrofossil leaf (*H. xylosteifolium* (Spach.) N. Rob son) from the Upper Pliocene of Georgia (Arbuzova, 2005); and a fossil seed of *H. danicum* Friis (Mai, 1995) along with other Pliocene *Hypericum* seed morphotypes found in Italy (Martinetto & al., 2006; Ciangherotti & al., 2007), Germany (Reid & Reid, 1915) and Belarus (Velich kevich & Zastawniak, 2003).

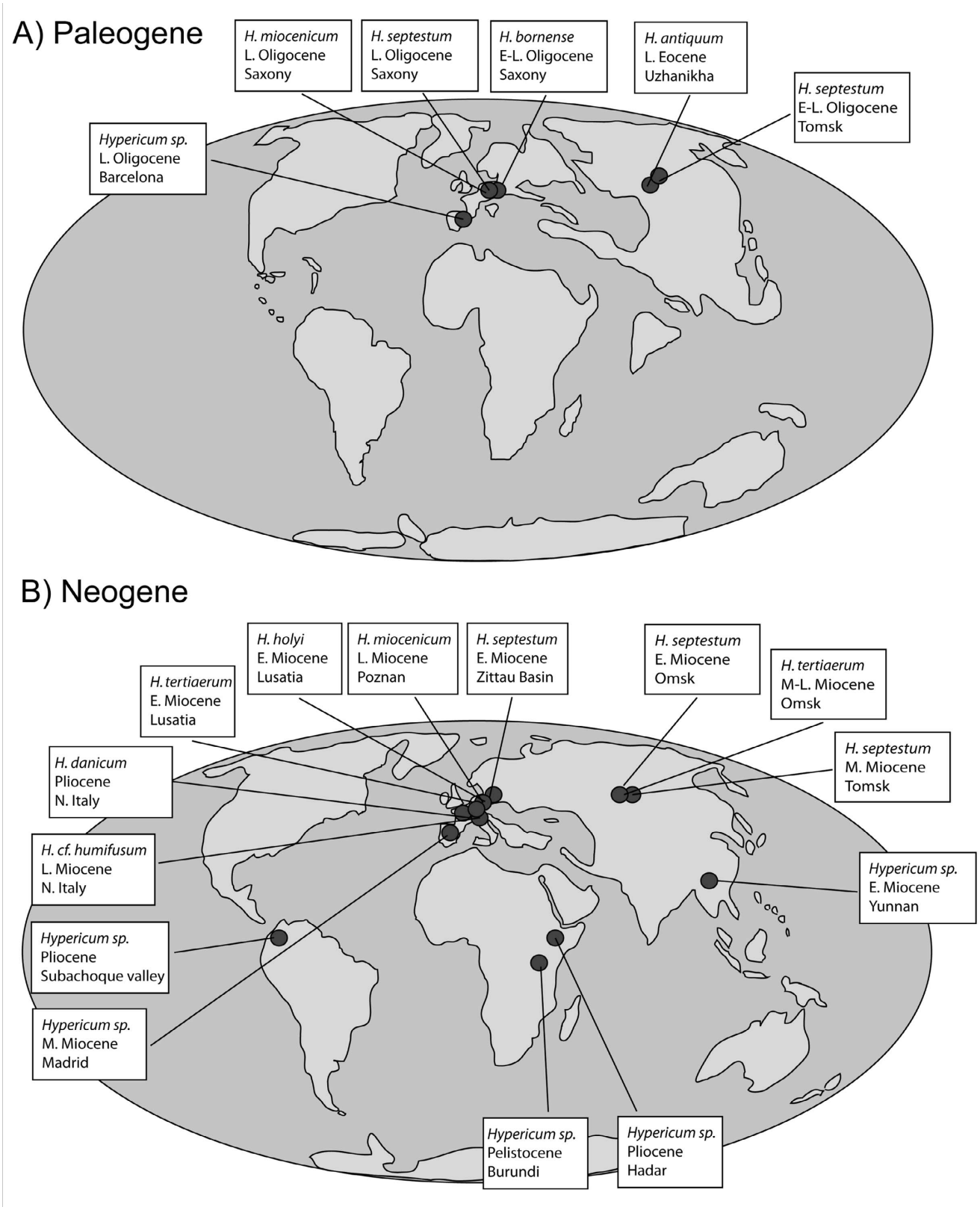


Fig. 4. Map showing the approximate distribution of the most relevant fossil remains of *Hypericum* discussed in relation to palaeogeographic re constructions: **A,** Paleogene (EoceneOligocene); **B,** Neogene (MiocenePliocene).

Apart from seeds and leaves, fossil pollen has been described from the Pliocene of diverse parts of the world, in cluding Colombia (Wijninga & Kuhry, 1989) and Ethiopia (Bonnefille & al., 1987). Pleistocene fossil pollen has been described from sites in the African Republic of Burundi (Bonnefille & Riollet, 1988; Bonnefille & al., 1992). How ever, Pleistocene pollen records are too abundant to review them in this work. If we take into account the amount of *Hypericum* citations in the Neogene fossil record, the most striking gap to our knowledge is the lack of fossil evidence available for *Hypericum* in North America until the Pleistocene: seeds of *Triadenum virginicum* (= *Hypericum virginicum*) represents the earliest record for the genus (Miller & Calkin, 1990).

Finally, we should comment the occurrence of some seed remains that after further revision have been excluded from the fossil record of *Hypericum* (Mai, 2001): *H. cf. ascyron* L.C. & E. M. Reid (1915), *H. cantalense* Reid (1923), *H. pliocenicum* Nikitin (1935) and *H. rostriferum* Jakubovskaja (1988).

5. Palaeogeographic implications

The palaeogeographical distribution of the main fossil remains of *Hypericum* discussed in this survey illustrates the evidence for the long history of this genus (Fig. 4). Interestingly, many of the abovementioned fossils have been as signed to species that are not presently distributed in the area in which the fossil was found. For example, remains attributed to *H. septestum* (described from Siberia, Europe) have been assigned to different extant species, none of which overlap in distribution with the fossil sites where the original species was described: e.g., *H. scabrum* L., is distributed from Lebanon to China, *H. aegypticum* L. occurs in northern Africa, and *Triadenum virginicum* (L.) Raf. (*H. virginicum* L.) is distributed in eastern North America. Another interesting example is the fossil remains attributed to *H.* cf. *balearicum* described from sites in Moldavia. The species under this name is now endemic to the Balearic Archipelago in the Western Mediterranean region. Many of these problematic fossils should be subject to careful revision before being assigned to extant species (Mai, 2001). It is also interesting to note that most fossil remains described here, especially the oldest records, are from sites in Central Europe and Russia, which could be a potential bias of the fossil record. For example, the lack of fossil record of *Hypericum* in North America is surprising since several species of *Hypericum* are endemic to this region. Despite this potential bias, we think that the distribution of the fossil record of *Hypericum* (see below) remains an important tool to gives us clues on its past distribution and biology.

The discovery of the fossil species *H. antiquum* in a site in Siberia dated from the Upper Eocene suggests that lineages of *Hypericum* were present in northern Eurasia during a period when the climate was considerably warmer and humid than today, following the Early Eocene Warming Event dated 55 million years ago (Wolfe, 1975; Tiffney, 1985a). At that time a "boreotropical" forest belt characterised by a mixture of deciduous hardwoods and evergreen subtropical elements, extended from Europe to North America and Asia, across a narrower Atlantic Ocean and a warmer Beringia (Wolfe, 1969, 1975, 1978; Tiffney 1985a, 1985b; Sanmartín & al., 2001). The global climate cooling that started with the Terminal Eocene Event at the end of the Eocene extirpated much of this forest from Siberia and other Northern Hemisphere regions, leaving extant remnants in eastern Asia and eastern North America (Tiffney, 1985a; Sanmartín & al., 2001). It seems plausible that the oldest lineages of *Hypericum* were part of this forest, or of its successor, the "mixed-mesophytic forest" (Tiffney, 1985a), as evidenced by the presence of fossil re mains *H. bornense* and *H. septestum* in Lower Oligocene sites of Germany and West Russia. These findings also suggest that the *Hypericum* species currently present in Siberia are derived from younger, more recently diverged lineages, while older Northern European lineages would have be come extinct during the glaciations of the Quaternary period (Sanmartín & al., 2001). Moreover, if our assumption is true, and *Hypericum* stem lineages were part of the Northern Hemisphere boreotropical forest belt, we might expect to find some Paleogene fossils in North America. A phylogeny-based biogeographic reconstruction that includes all these fossil remains is needed to confirm the importance of these findings and our paleobiological conclusions (Meseguer & al., in prep.).

Bibliography

Arbuzova, O. 2005. *Hypericum* L. In: Budantsev, L. (eds.), *Iskopaemye tsvetkovye rastenija Rossii i sopredel'nyh gosudarstv [Fossil flowering plants of Russia and adjacent countries]*, *Vol. 4 Nyctaginaceae-Salicaceae*. Izdatelstvo Nauka Leningradskoe otdnie, 1974. Leningrad. (In Russian).

Arteaga, L.L. 2007. *Vismia glaziovii* Ruhl. (Guttiferae) seed size and its re lationship with the germination speed and seedling size. *Revista Peruana de Biología* 14: 1720.

Baranowska-Zarzycka, Z. 1988. Main features of the Pliocene fruit-seed flora from Ruszów near Z˙ary (West Poland). *Acta Palaeobotanica* 28: 2327.

Barrón, E., RivasCarballo, R., Postigo-Mijarra, J.M., Alcalde-Olivares, C., Vieira, M., Castro, L., Pais, J. & Valle-Hernández, M. 2010. The Ceno zoic vegetation of the Iberian Peninsula: A synthesis. *Review of Palaeobotany and Palynology* 162: 382402 (SI).

Barros, M. & Ramos, A. 1984. *Estudio del polen de las especies de Hyperi cum sect. Hirtella Stef. en la Península Ibérica y Baleares*. V Simposio de Palinología. APLE. Resúmenes. Córdoba.

Bonnefille, R., Vincens, A. & Buchet, G. 1987. Palynology, stratigraphy and palaeoenvironment of a Pliocene hominid site (2.9–3.3 M.Y.) at Hadar, Ethiopia. *Palaeogeography, Palaeoclimatology, Palaeoecolog* 60: 249281.

Bonnefille, R. & Riollet, G. 1988. The Kashiru pollen sequence (Burundi). Palaeoclimatic implications for the last 40000 yr B.P. in tropical Africa. *Quaternary Research* 30: 1935.

Bonnefille, R., Chali, F., Guiot, J. & Vincens, A. 1992. Quantitative esti mates of full glacial temperatures in equatorial Africa from palynologi cal data. *Climate Dynamics* 6: 251257.

Cavagnetto, C. & Anadón, P. 1996. Preliminary palynological data on floristic and climatic changes during the Middle EoceneEarly Oligo cene of the eastern Ebro Basin, northeast Spain. *Review of Palaeo botany and Palynology* 92: 281305.

Crepet, W.L. & Nixon, K.C. 1998 . Fossil Clusiaceae from the late Creta ceous (Turonian) of New Jersey and implications regarding the history of bee pollination. *American Journal of Botany* 85: 11221133.

Ciangherotti, A., Esu, D., Martinetto, E. & Giuntelli, P. 2007. The re markable Middle Pliocene nonmarine mollusc record from Ceresole d'Alba, Piedmont, northwest Italy: Biochronology, palaeobiogeogra phy and palaeoecology supported by fossil plants. *Geobios* 40: 573587.

Clarke, G.C.S. 1975. Irregular pollen grains in some *Hypericum* species. *Grana* 15: 117125.

Clarke, G.C.S. 1976. The northwest European pollen flora, 7 Guttiferae. *Review of Paleobotany and Palynology* 21: 125142.

Clarke, G.C.S. 1981. 4. Pollen morphology. In: Robson, N.K.B., Studies in the genus Hypericum L. (Guttiferae) 2. Characters of the genus. *Bul letin of the British Museum of Natural History (Botany)* 8: 115118.

Corner, E.J.H. 1976a. *The Seeds of Dicotyledons*, Vol. 1. Cambridge Uni versity Press, Cambridge.

Corner, E.J.H. 1976b. *The Seeds of Dicotyledons*, Vol. 2. Cambridge Uni versity Press, Cambridge.

Davis, C.C., Webb, C.O., Wurdack, K.J., Jaramillo, C.A. & Donoghue, M.J. 2005. Explosive radiation of Malpighiales supports a midCreta ceous origin of modern tropical rain forests. *Amer. Naturalist* 165: E36–E65.

Dorofeev, P.I. 1957. Novye dannye o pliotsenovoy flore Kamy (New data on the Pliocene flora of Kama). *Doklady Akademii Nauk SSSR* 117: 487490. (In Russian).

Dorofeev, P.I. 1959. *Materialyk poznaniyu miotsenovoy flory Rostovskoy oblasti (summary: Contribution to the study of Miocene floras of the Ros tov region). Problemy botaniki (The problems of Botany)* 4: 143189. (In Russian).

Dorofeev, P.I. 1963. *Tretichnye flory Zapadnoi Sibiri (Tertiary Floras of Western Siberia)*. Izdatelstvo Akademii Nauk, Moscow/Leningrad. 343 pp.

Dorofeev, P.I. 1979. O pliotsenovoy flore s. Dan'shino na Donu (On the Pliocene flora of the vill. Dan'shino on the Don): 87–94. In: Raskatov G.I. (eds.), *Problemy Antropogena tsentralnykh rayonov Russkoy plat formy (The Problems of the Antropogene of central districts of the Russian platform)*. Izdatelstvo Voronezh University, Voronezh. (In Russian).

Dorofeev, P.I. 1986. *Iskopaemye Potamogeton (Fossil Potamogeton)*. Izda telstvo Nauka, Leningrado. (In Russian).

Dyjor, S., Kvaček, Z., ŁańcuckaŚ rodoniowa, M., Pyszyński, W., Sadowska, A. & Zastawniak, E. 1992. The younger Tertiary deposits in the Gozdnica region (SW Poland) in the light of recent palaeobotanical re search. *Polish Botanical Studies* 3: 1129.

Friis, E.M., 1985. Angiosperm fruits and seeds from the Middle Miocene of Jutland (Denmark). *Biologiske Skrifter, Det Kongelige Danske Vi denskabernes Selskab* 24: 1165.

Gümbell, F. & Mai, D.H. 2002. Neue Pflanzenfunde aus dem Tertiär der Rhön. Teil 1: Miozäne Fundstellen. *Mitteilungen aus dem Museum für Naturkunde in Berlin, Geowissenschaftliche Reihe* 5: 345384.

Gümbell, F. & Mai, D.H. 2004. Neue Pflanzenfunde aus dem Tertiär der Rhon. Teil 2: Pliozäne Fundstellen. *Mitteilungen aus dem Museum für Naturkunde in Berlin, Geowissenschaftliche Reihe* 7: 175220.

Hao, H., Ferguson, D.K., Feng, GP., Ablaev, A., Wang, ·YF. & Li, CS.2010. Early Paleocene vegetation and climate in Jiayin, NE China. *Climatic Change* 99: 547566.

Ho, S.Y.W. 2007. Calibrating molecular estimates of substitution rates and divergence times in birds. *Journal of Avian Biology* 38: 409414.

Holý, F. 1974. *Neogénní mastixioidní květena svrchního slojového pásma z lomu Kristina (Hrádek n. N.)*. Dissertation thesis, Department of Palaoentology, National Museum, Prague. 131pp.Jahn, A., ŁańcuckaŚ rodoniowa, M. & Sadowska, A. 1984. Stanowisko ut worów plioceńskich w Kotlinie Kłodzkiej (summary: The site of Pliocene deposits in the Kłodzko Basin, Central Sudetes). *Geologia Sudetica* 18: 743.

Khan, H.A. 1969. Pollen morphology of Indian Hypericaceae. *Journal of Palynology*. 5: 9799.

KovarEder, J., Kvaček, Z., Martinetto, E. & Roiron, P. 2006. Late Miocene to Early Pliocene vegetation of southern Europe (74 Ma) as reflected in the megafossil plant record. *Palaeogeography, Palaeoclima tology, Palaeoecology* 238: 321339.

Kvaček, Z. & Teodoridis, V. 2003. Tertiary macrofloras of the Bohemian Massif: a review with correlations within Boreal and Central Europe. *Bulletin of Geosciences* 82: 383408.

ŁańcuckaŚrodoniowa, M. 1981. Macroscopic plant remains from the Miocene deposits at Stara Wieśnear Wilamowice (southern Poland). *Acta Palaeobotanica* 21: 115126.

Mai, D.H. 1995. Palaeocarpological investigations in the Villafranchian (Pliocene) of Italy. *Bollettino Museo Regionale di Scienze Naturali* 13: 407437.

Mai, D.H. 1997. Die Oberoligozänen Floren am Nordrand der Sächsi schen Lausitz. *Palaeontographica B* 244: 1124.

Mai, D.H. 1999. Die untermiozänen Floren aus der Spremberger Folge und dem 2. Flözhorizont in der Lausitz. Teil I: Farnpflanzen, Konife ren und Monokotyledonen. *Palaeontographica B* 250: 176.

Mai, D.H. 2000. Die untermiozänen Floren aus der Spremberger Folge und dem 2. Flözhorizont in der Lausitz. Teil III: Dialypetale und Sym petale. *Palaeontographica B* 253: 1106.

Mai, D.H. 2001. Die mittelmiozänen und obermiozänen Floren aus der Meuroer und Raunoer Folge in der Lausitz. III. Fundstellen und Pala eobiologie. *Palaeontographica B* 258:185.

Mai, D.H. & Walther, H. 1978. Die Floren der Haselbacher Serie im Weisselster Becken (Bezirk Leipzig, DDR). *Abhandlungen des Staatli chen Museums für Mineralogie und Geologie zu Dresden* 28: 1200.

Mai, D.H. & Walther, H. 1988. Die pliozänen Floren von Thüringen. *Quartärpaläontologie* 7: 55295.

Martinetto, E., Scardia, G. & Varrone, D. 2007. Magnetobiostratigraphy of the stura di lanzo fossil forest succession (Piedmont, Italy). *Revista Italiana di Paleontologia e Stratigrafia* 119: 109125.

Mártonfi, P., Janíková, M. & Zezula, I. 2002. Palynological analysis of sev en *Hypericum* taxa. *Biologia, Bratislava* 57: 455460.

Meller, B. & Hofmann, C.C. 2004. Paleoecology of Diaspore and Paly nomorph assemblages form Late Miocene lake sediments (Mataschen near Fehring, East Styria, Austria). *Joannea Geologie und Paläontolo gie* 5: 177217. (In German).

Miller, N.G. & Calkin, P.E. 1992. Paleoecological interpretation and age of an interstadial lake bed in western New York. *Quaternary Research* 31: 7588.

Mourão, K.S. & Beltrati, C. 2001. Morphology and anatomy of developing fruits and seeds of *Vismia guianensis* (Aubl.) Choisy (Clusiaceae). *Re vista Brasileira de Biologia* 61: 147158.

Matzk, F., Meister, A., Brutovska, R. & Schubert, I. 2001. Reconstruction of reproductive diversity in *Hypericum perforatum* L. opens novel strategies to manage apomixis. *The Plant Journal* 26: 275282.

Negru, A.G. 1972. *Rannesarmatskaya flora severovostoka Moldavii (The Early Sarmatian flora from NorthEast of Moldavia)*. Shtiintsa, Kishinev. 169 pp. (In Russian).

Nikitin, P.A. 1948. Pliocenovye flory s reki Obi v rayonie Tomska (Pliocene floras from Ob river near Tomsk). *Doklady Akademii Nauk SSSR* 61: 11031106. (In Russian).

Nikitin, P.A. 1957. *Pliotsenovye i chetvertichnye flory Voronezhskoy oblasti (Pliocene and Quaternary floras of Voronezh district)*. Izdatelstvo Akademii Nauk SSSR, MoskvaLeningrad. (In Russian).

Nikitin, V.P. 2006. *Paleocarpology and stratigraphy of Paleogene and Neo gene strata of Asian Russia*. Izdatelstvo Akademii "Geo", Novosibirsk. (In Russian).

Nürk, N.M. & Blattner, F.R., 2010. Cladistic analysis of morphological characters in *Hypericum* (Hypericaceae). *Taxon* 59: 14951507.

Ohlendorf, O. 1907. *Beiträge zur Anatomie und Biologie der Früchte und Samen einheimischer Wasser und Sumpfpflanzen*. Inaugural Dissertati on, FriedrichAlexanderUniversität, Erlangen.

Palamarev, E., Bozukov, V., Uzunova, K., Petkova, A. & Kitanov, G. 2005. Catalogue of the Cenozoic plants of Bulgaria (Eocene to Plio cene). *Phytologia Balcanica* 11: 215364.

Ramos Núñez, A. 1983. Estudio biosistemático del género *Hypericum* L. (Guttiferae) en la Península Ibérica e Islas Baleares. 1. Caracteres se minales. *Trabajo del Departamento de Botánica* 12: 4562.

Reid, C. & Reid, E.M. 1915. The Pliocene floras of the DutchPrussian border. *Mededeelingen van de Rijksopsporing van Delfstoffen* 6: 1178. Reid, E.M. 1923. Nouvelles recherches sur les graines du Pliocéne infe rieur du PontduGail (Cantal). *Le Bulletin de la Société géologique de France* IV, 23: 305355.

Robson, N.K.B., 1977. Studies in the genus *Hypericum* L. (Guttiferae). 1. Infrageneric classification. *Bulletin of the British Museum of Natural History (Botany)* 5: 295355.

Robson, N.K.B. 1981. Studies in the genus *Hypericum* L. (Guttiferae). 2. Characters of the genus. *Bulletin of the British Museum of Natural His tory (Botany)* 8: 55226.

Ruhfel, B.R., Bittrich, V., Bove, C.P., Gustafsson, M.H.G., Philbrick, C.T., Rutishauser, R., Xi, Z & Davis, C.C. 2011. Phylogeny of the clu sioid clade (Malpighiales): evidence from the plastid and mitochondri al genomes. *American Journal of Botany* 98: 306325.

Sanmartín, I., Enghoff, H. & Ronquist, F. 2001. Patterns of animal disper sal, vicariance and diversification in the Holarctic. *Biological Journal of the Linnean Society* 73: 345390.

Stevens, P.F. 2007. *Hypericaceae*. In: Kubitzki, K. (eds.), *The families and genera of vascular plants*, Vol. 9. Pp. 194201. Springer, Berlin, Heidel berg.

Teodoridis, V. 2002. Notice about a revision of the Early Miocene carpo logical material from the Czech part of the Zittau Basin. *Zprávy o geo lofických výzkumech v roce*: 157158.

Teodoridis, V. 2003, Early Miocene carpological material from the Czech part of the Zittau Bassin. *Acta Palaeobotanica* 43: 949.

Thomas, J.L. 1970. Haploid a diploid pollen in *Hypericum patulum*. *Jour nal of the Arnold Arboretum* 51: 247250.

Tiffney, B.H. 1985a. Perspectives on the origin of the floristic similarity between eastern Asia and eastern America. *Journal of the Arnold Ar boretum* 66: 7394.

Tiffney, B.H. 1985b. The Eocene North Atlantic land bridge: its impor tance in Tertiary and modern phytogeography of the Northern Hemi sphere. *Journal of the Arnold Arboretum* 66: 243273.

Van der Burgh, J. 1987. Miocene floras in the lower Rhenish basin and their ecological interpretation. *Review of Palaeobotany and Palynology* 52: 299366.

Velichkevich, F.Yu. 1982. Pleystotsenovye flory lednikovykh oblastey Vos tochnoEvropeyskoy ravniny (The Pleistocene floral of glacial areas of the East European Plain). *Nauka i Tekhnika, Minsk*. (In Russian).

Velichkevich, F.Yu. 1990. Pozdnepliotsenovaya flora Dvortsa na Dnepre (The Late Pliocene flora of Dvorets on the Dnieper River). *Nauka i Tekhnika, Minsk*. (In Russian).

Velichkevich, F.Yu. & Zastawniak, E. 2003. The Pliocene flora of Kholmech, southeastern Belarus and its correlation with other Pliocene floras of Europe. *Acta Palaeobotanica* 43: 137259.

Wijninga, V.M. & Kuhry, P. 1990. A Pliocene flora from the Subachoque Valley (Cordillera Oriental, Colombia). *Review of Palaeobotany and Palynology* 62: 249290.

Wolfe, J.A. 1969. Neogene floristic and vegetational history of the Pacific Northwest. *Madroño* 20: 83110.

Wolfe, J.A. 1972. An interpretation of Alaskan Tertiary floras. In: Gra ham, A. (eds.), *Floristics and Paleofloristics of Asia and eastern North America:* 201233. Elsevier, Amsterdam.

Wolfe, J.A. 1975. Some aspects of plant geography of the northern hemi sphere during the Late cretaceous and Tertiary. *Annals of the Missouri Botanical Garden* 62: 264279.

Yakubovskaya, T.V. 1982. Pliotsenovye flory Belorusskovo Podneprovýa (Pliocene floras of Byelorussian Podnieprovye): 3461. In: Velichke vich, F.Yu. (eds.) *Paleokarpologicheskie issledovaniya kaynozoya (Pa laeocarpological investigations of Cenozoic)*. Nauka i Tekhnika, Minsk. (In Russian).

Yakubovskaya, T.V. 1984. *Ocherk neogena i rannevo antropogena Pone manýa (Sketch of Neogene and Early Antropogene of the Ponemanye).* Nauka i Tekhnika, Minsk. (In Russian).

Zhaoa, L.C. , Wanga, Y.F., Liua, C.J. & Li, C.S. 2004. Climatic implica tions of fruit and seed assemblage from Miocene of Yunnan, south western China. *Quaternary International* 117: 8189.

CHAPTER 4

Using fossils to reveal the impact of ancient climate change in plant evolution

This chapter has been done in collaboration with: Jorge M. Lobo (Museo Nacional de Ciencias Naturales, CSIC, Madrid), Richard Ree (Field Museum of Natural History, Chicago), David J. Beerling (Department of Animal and Plant Sciences, University of Sheefield, United Kingdom) and Isabel Sanmartín (Department of Biodiversity and Conservation, Real Jardín Bótanico-CSIC, Spain).

This chapter has been submitted to *PLoS Biology*.

CHAPTER 4

Using fossils to reveal the impact of ancient climate change in plant evolution

Andrea Sánchez Meseguer, Jorge M. Lobo, Richard Ree, David J. Beerling, and Isabel Sanmartín

Abstract

The last sixty million years have been characterized by numerous and large climatic oscillations. Inferring how organisms responded to them, by adaptation, extinction, or migration, has important implications for understanding the effects of global warming on Earth's biodiversity. Previous approaches to this question have been hampered by the fact that the biodiversity we observe today is only a small fraction of that of the past. Fossils can help bridge this gap by providing information on when, where, and in which climatic conditions a lineage's ancestors lived. Here, we use a fossil-informed approach to niche modeling and biogeographic analysis to understand the role of ancient climate change in driving the evolutionary history of genus *Hypericum*, one of few plant genera hypothesized to have managed the transition from tropical to temperate climates. Our results reveal that Hypericum stem-lineages were already distributed in the Holarctic before the Late Eocene crown-group diversification and that niche conservatism has been prevalent over the Cenozoic evolutionary history of the genus. Early lineages had probably wider climatic tolerances than their more recent descendants, and divergence among major clades was driven by niche specialization within these ancestral conditions rather than by a change to a new ecospace. Geographical expansion in the Late Cenozoic was mediated by the existence of climatic corridors, such as Beringia or the North African-Arabian bridge, allowing lineages to track their niches over space and time. Conversely, the latest Quaternary climatic fluctuations led to a shrinking in geographic distribution and the continental disjunctions we observe today.

1. Introduction

Fifty-five million years ago, the Earth experienced one of the warmest intervals in Earth history, with global temperatures up to 10 degrees warmer than present and tropical climates characterizing northern latitudes ([1]). This "green- house" world ([6]) came abruptly to an end at the Terminal Eocene Event (TEE, 35 Ma, [1]), when a sharp decline in atmospheric CO2 levels and a change in continental configuration brought about a dramatic drop in global temperatures. The period that followed, the Neogene, was characterized by rapid climatic oscillations between warm and humid and dry and cold episodes ([1]). Since the Quaternary, an "icehouse" world with long glacial phases and short warm humid inter- glacial periods, has characterized the Earth (Fig. 1), but human activity and continuing increases in atmospheric carbon dioxide (CO_2) levels ([6]), threaten to bring about conditions that resemble a "greenhouse" world. Global warming and its effects in the current biodiversity has attracted renewed attention to these events of ancient climate change. Under- standing how organisms responded to past climatic events might help predict how they will cope with current and future climate changes ([5]). Plants are strongly tied to climatic conditions, making them ideal candidates to evaluate the impact of past environmental change on evolutionary trajectories over geological time scales ([5]). Paleobotanical evidence (Fig. 1) shows that in the face of rapid climate

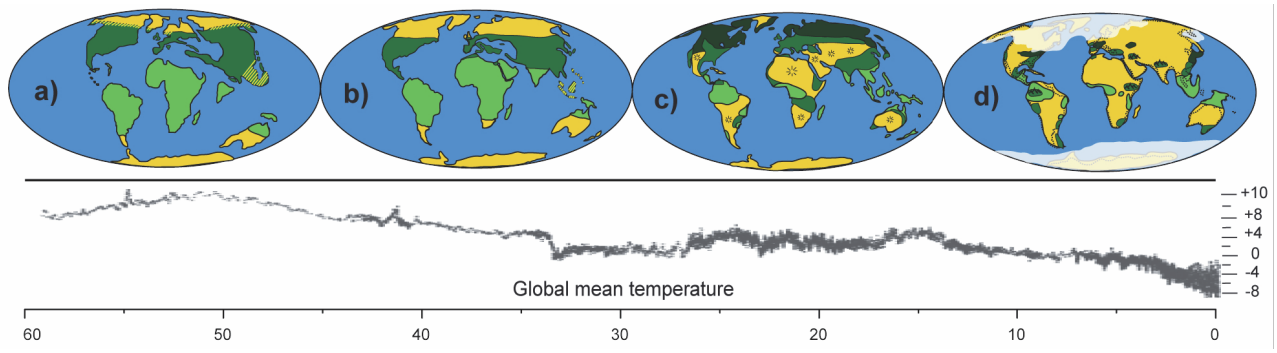


Fig. 1. Scheme representing changes in vegetation composition in response to major shifts in Cenozoic climate. a) During the warm Early Eocene, a "boreotropical" forest, composed by hardwood deciduous and broad-leaved taxa, extended across North America and Eurasia ([2], [3], [4] [5]). b) After the TEE event, this forest was replaced by a temperate, deciduous mixed-mesophytic forest, whereas evergreen boreotropical elements migrated southwards or became extinct. c) Global climate cooling starting in the Late Miocene led to the expansion of a boreal coniferous forest across northern Eurasia and North America, displacing the temperate forest ([7], [8]). d) During the Quaternary, ice sheets covered large regions of Eurasia and

change, plants have responded through either adaptation (change of niche by evolving new traits), geographic movement (tracking their preferred niche by migration), or extinction (biotic loss) ([2], [3], [5]). Niche conservatism ([9] seems to be prevalent in plants ([10], [11]). For example, the excess of species in tropical regions compared with temperate latitudes has been explained by the fact that many plant families originated in the Late Cretaceous-Early Cenozoic tropical conditions and preserved these ancestral climatic preferences over evolutionary time, while only few of these lineages developed the relevant traits to adapt to the more seasonal temperate climates ([12], [11]).

Tracing the response of organisms to deep-time climate change requires the integration of several disciplines traditionally separated by different theoretical and methodological approaches, such as phylogenetics, ecology, and paleontology ([13], [9], [14]). Recent studies have explored combining Eco- logical Niche Models (ENM) with molecular phylogenies to trace changes in climate niche variables and identify cases of niche evolution ([15], [16]), or to detect areas that were climatically adequate in the past ([17]). These approaches, how- ever, are limited by the fact that they are based eminently on present-day distributions and assume to some extent that ancestral lineages shared the same ecological requirements than their extant descendants, which might be true for short time scales but less likely over geological scales of millions of years ([29], [14]). Present observations represent but the small diversity fraction that survived to the present, which results in more uncertain ancestral inferences as we move to the past along the phylogeny. The fossil record can help us bridge this gap by providing information on when, where, and in which climatic conditions the ancestors of a lineage lived. Fossils have traditionally been used in biogeography to provide calibration points for phylogenetic dating, but recent studies have shown that incorporating the spatial range of fossils can change dramatically the inferred biogeographic scenario ([18], [19]). These studies, however, have relied on using the fossil as an additional lineage in the phylogeny, which requires coding the morphological traits of the fossil taxa alongside the extant molecular characters ([18]), or assuming arbitrary branch lengths for the fossil lineage ([19]). On the other hand, paleoecologists have used ENM models based on fossils to re- construct ancestral niches and track patterns of niche conservatism versus evolution over longer time scales ([20], [14]). So far, the incompleteness of the fossil record and the lack of environmental data in the deep past has limited this approach to reduced geographical regions ([20]) or the recent geological time ([21]).

Here, we used a fossil-informed approach to ecological niche modeling and biogeographic reconstruction to under- stand the role of Earth's climate change in

driving the evolution of *Hypericum*, St. John's wort, one of 100 largest genera of angiosperms (ca. 500 species). *Hypericum* has a worldwide distribution, a rich fossil record dating back to the Early Cenozoic ([22]), and well-documented phylogenetic relationships ([23], [24]), which makes it an ideal group to explore the potential of fossils to place ancestors in space and time in a deep-time phylogenetic context. Furthermore, *Hypericum* has been hypothesized as one of few plant taxa that succeeded in the transition from tropical to temperate climates ([11]). The highest diversity of *Hypericum* is currently found in the temperate regions of the Northern Hemisphere, while most of its relatives, members of the "clusioid clade" from the order Malpighiales, are tropical forest woody plants ([25], [26]). A recent biogeographic study reconstructed *Hypericum* ancestors as Late Cretaceous tropical plants that migrated from Africa into the Western Palearctic coincident with the expansion of the boreotropical forest in the Early Tertiary ([23]). After the TEE, *Hypericum* dispersed to other Holarctic continents, while presumably developing adaptations to colder cli- mates, such as the herbaceous habit or unspecialized corollas ([23]. However, this analysis was based only on the distribution of extant taxa and did not consider the information contained in the fossil record. The earliest fossil representative of the genus, the Late Eocene *H. antiquum* from Siberia ([22]), comes from a region not inferred in previous biogeographic reconstructions and that it is now at the fringe of the genus distribution, but was in the Late Eocene covered by the boreotropical forest (Fig. 1a). Here, we present a new approach to integrate the information provided by fossils, both spatial (ancestral ranges) and ecological (climatic preferences of extinct lineages), into biogeographic analysis in order to explore the evolutionary response of plants to deep-time climatic events. First, we constrained the inference of ancestral ranges for those nodes in the phylogeny calibrated with the fossil record to include the geographic distribution of the fossil. Second, we use recent coupled ocean-atmosphere global climate model simulations that extend into early Cenozoic ([27], [6], [28]) to reconstruct the environment of present and extinct lineages and to examine the role of niche conservatism versus niche evolution in explaining *Hypericum* present diversity patterns. Finally, we integrated these fossil ranges and climate niche models into biogeographic analysis to detect regions that were in the past within the climatic tolerance range of *Hypericum*, and might have acted as migration corridors allowing species to track their climatic preferences through space and time.

2. Results

2.1. Ecological Niche Modeling.

We estimated the climate niche optimum and tolerance range of *Hypericum* based on either current distributions (extant *Hypericum*) or fossil record data ("greenhouse" and "coldhouse" ancestral *Hypericum* lineages), and projected these optima onto six different paleoclimate scenarios representing major shifts in Cenozoic cli- mate. Although the percentage of climatically adequate area was lower in fossil-based projections – probably due to the small number of data points – the overall temporal pattern was similar between extant and ancestral lineages: the warm Early Eocene simulation generated the lowest number of world cells within the climate optimum of *Hypericum* (blue color in maps), and the Pliocene simulation the highest number (Fig. 2a,b; Supplementary Information Figures S1–S2). Also, most greenhouse and coldhouse fossil sites were located within cells with high values of climatic adequacy according to the suitability derived from extant *Hypericum* lineages (Fig. S3). Projections of the greenhouse-fossil optimum generated favor- able conditions in all time layers, including the most recent simulations, whereas coldhouse lineages showed slightly worse climatic adequacy for the early warmer "climate worlds" (Fig. S2); similar results were obtained after removing the inter- mediate Miocene fossils (Fig. S4). During the warm and tropical Early Eocene period (Fig. 2a),

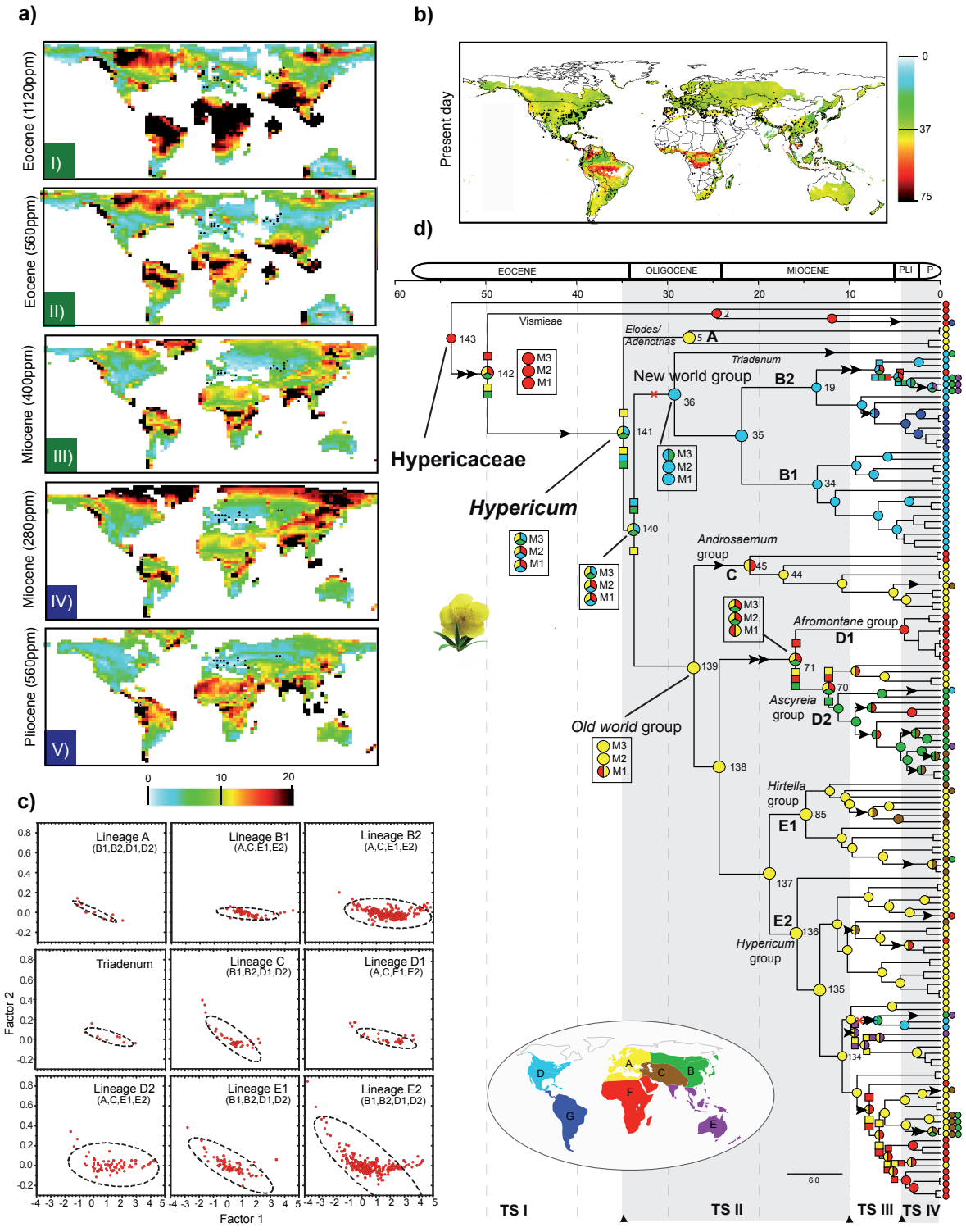


Fig. 2. Integrative approach to reconstruct the spatio-temporal evolution of extant and extinct lineages using phylogenetic, ecological, and fossil-record information. A) Climatic preferences of "ancestral" *Hypericum* lineages projected onto six paleoclimate scenarios and based on two sets of fossils: Eocene-to-Miocene ("greenhouse") fossils (I–III) and Late Miocene-to-Pliocene ("coldhouse") fossils (IV–V); dots represent fossil locations. B) Climatic preferences of "extant" *Hypericum* species based on current distributions and projected onto present-day climate data. C) Niche segregation along ENFA Factors 1 ("tropicality") and 2 ('aridity") for major phylogenetic lineages within *Hypericum* (see D); lineages showing significant differences with the one in the figure are given between brackets. D) Biogeographic reconstruction showing the effects of adding new sources of information to the analysis: M1 uses only current distributions; M2 adds a paleogeographic model to reflect continental connectivity through time (time slices (TS) indicated by grey vertical bars); M3 incorporates fossil ranges; M4 (main figure) further integrates estimations of ancestral distributions and predicted climatic migration corridors (Fig. 2a). Node ancestral ranges (pie charts), range division scenarios (square splits), and dispersal (arrows) and extinction (red crosses) events, were inferred using DEC ([42]) over the dated phylogeny of Meseguer et al. 2013 ([23]) and 1000 trees from the posterior distribution (see SI Text for more details). Colour codes for nodes and terminal taxa correspond with those in the inset map.

only the northernmost areas showed adequate climatic conditions for *Hypericum* greenhouse ancestors, whereas the cooling event at the end of the Eocene created favourable conditions for these lineages in the intermediate latitudes of North America and Eurasia, including Beringia (Fig. 2a–II). As the climate became gradually cooler from the Miocene onwards, the geographical area with favourable conditions increased, especially in the Pliocene (Fig. 2a–III-V), while the Preindustrial world marked a general reduction in geographic distribution for coldhouse lineages (Fig. S2). We also analyzed extant climatic preferences among major phylogenetic clades in *Hypericum*. An ENFA analysis (see Methods) gave us two composite climatic variables with the highest power to explain the distribution of *Hypericum*: "tropicality" and "aridity" (Fig. S5). The first factor segregated the "more tropical" American and Asian-African clades B and D (the "New World group" and *Ascyreia*) from the mainly Eurasian clades A, C, and E, which exhibit less tropical affinities (Fig. 2c). No significant differences were found along the "aridity" factor, i.e., all extant *Hypericum* species live outside deserts and humid tropical regions (Fig. 2b,c); similar constraints were observed in the ancestral climate projections of fossil lineages (Fig. 2a).

2.2. Biogeographic analyses.

All four biogeographic models reconstructed the ancestor of Hypericaceae (node 143, Fig. 2d, Table S1) as African, but the timing of the first migration event to the Holarctic changed depending on the evidence included. Models based on spatial distributions, either extant (M1–M2) or fossil-extant (M3), inferred migration to the Holarctic along the branch leading to crown-lineage *Hypericum*, ca. 35 Ma (node 141, Figs. S6–S8), whereas M4, which integrates also niche model predictions, reconstructed the stem-lineage of *Hypericum* (node 142) as already present in the Holarctic by the Early Eocene (ca. 50 Ma), with migration from Africa along the branch leading to the ancestor of Hypericeae-Vismieae (Fig. 2a, Fig. S9, Table S2). Another difference between models concerns the colonization of the Eastern Palearctic region. Models based only on extant distributions (M1–M2) reconstructed the ancestor of *Hypericum* as Western Palearctic, Nearctic and Africa (node 141, Fig. 2d), whereas those including the geographic range of the fossil *H. antiquum* (M3–M4) inferred crown *Hypericum* to be widespread in the entire Palearctic and Nearctic regions, with prior extinction in Africa (Fig. 2d). Fossil-based M3 model reconstructed the New World group as widespread in the Eastern Palearctic, while M4 inferred an extinction event in this region before the first diversification event, ca. 35–30 Ma (node 36, Fig. 2d). All models, except M1, inferred a new colonization of the Eastern Palearctic in the Early Miocene (ca. 16 Ma) along the branch leading to the *Ascyreia* clade (node 71, Fig. 2a).

3. Discussion

3.1. Global niche conservatism and lineage divergence.

Because evolutionary shifts from one ecological niche into another generally require substantial physiological adjustments i.e., the evolution of frost tolerance, Donoghue ([11]) argued that it is often easier for plant lineages to move into regions with the same climatic conditions than to adapt to new ones. *Hypericum* is the largest genus in the clusioid clade and the only one that has successfully diversified in temperate areas and achieved a worldwide distribution. This success has been hypothesized to be related to a change in climatic preferences early in the history of *Hypericum* that allowed it to adapt to the colder environments of the northern latitudes ([23]). Since the fossil record of the genus (and Hypericaceae) does not extend further back than the Late Eocene ([22]), we cannot rule out whether *Hypericum* ancestors were already preadapted or evolved temperate tolerances *in situ*, concurrent with the crown-group diversification and the TEE event. Davis et al. ([25]) reconstructed the ancestors of Hypericaceae as inhabitants of open woodland areas in tropical latitudes, and, al- though their analysis was based on only two genera, it sug- gests the possibility that *Hypericum* stem-lineages were already adapted to more seasonal conditions before the

first divergence of the genus. Instead, what our climate projections suggest is that the present tolerances of *Hypericum* were already achieved by the Late Eocene, coincident with the crown-group diversification, and that these preferences have remained stable over its entire Cenozoic history (Fig. 2). We did not observe any significant shift in the suitable climatic conditions derived from greenhouse or coldhouse fossil lineages, or between the latter and living *Hypericum*. The predicted distribution for *Hypericum* across time based either on fossil or extant tolerances was to a great extent the same. Despite several range contractions and a late expansion to the Southern Hemisphere (Fig. 2a), *Hypericum* ancestors inhabited similar climatic regions in the Holarctic than their extant descendants, and could not live on deserts or tropical regions. This degree of "niche conservatism" seems surprising, and suggests that climatic tolerances are preserved even over millions of years ([29]).

Willis and MacDonald ([5]) argued that many tropical species that evolved during the Late Cretaceous in much higher temperatures and CO2 levels developed a wider climatic tolerance than is apparent from present day distributions, which allowed them later to survive through the rapid climate changes of the Cenozoic. Over the course of evolution this genetic resilience would have been lost across genetic splits, so that descendant lineages "specialized" in a different subset of the original niche. Although the climatic niche of *Hypericum* as a whole seems to have been conserved over its entire evolutionary history (Fig. 2a,b), we detected climatic differences among extant major clades within the genus (Fig. 2c). Also, a slight reduction in niche can be observed between the early greenhouse fossil projections, with climatically adequate areas in all paleosimulations, and the more recent coldhouse projections, which showed poor adequacy to the early worlds (Figs. S2–S3). We recognize that the number of fossils studied is not sufficient to confirm this hypothesis. However, the general picture that emerges is one in which *Hypericum* ancestral lineages, probably descendants of Late Cretaceous tropical plants ([25], already had the plasticity to cope with the large temporal and spatial variability of Cenozoic climates, early in its evolutionary history. Lineage divergence within the genus then being driven by niche specialization within these ancestral conditions rather than by a shift to a new ecospace ([14]). Interestingly, the percentage of geographic space that was climatically adequate for fossil and extant *Hypericum* was larger in the Pliocene than in the Preindustrial and Present worlds (Fig. 2a,b, Figs S1–S3). As with many other plant taxa that succeeded in the transition to temperate habitats, the latest events of Pleistocene climate cooling were probably too extreme for *Hypericum*. Once an ecological threshold was passed, species could not persist and retreated or reduced their geographic range ([11]), which can be observed in the disjunct distribution of some species today, with sister groups at both sides of Beringia, e.g., *Triadenum* (Fig. 2d).

3.2. Fossils place ancestors in time, space, and climate.

The importance of fossils in biogeography has been compared to the problem of incomplete taxon sampling in phylogenetics ([32]). Failing to include all representatives within a clade might lead to biased reconstructions of divergence times ([30]) or evolutionary history ([31]). Extinction causes a similar effect in biogeography. Integrating the geographical range of extinct lineages often reveals unexpected biogeographic patterns that change our view of a lineage's evolution ([32], [18]). Incorporating the range of the oldest fossil record in *Hypericum* resulted in the Eastern Palearctic region being inferred as part of the ancestral distribution (Fig. 2d), instead of a late colonization event ([23]). Also, our fossil-based projections showed climatically suitable areas in this region during the Early Cenozoic (Fig. 2a). This disagreement between extant biogeographic patterns and fossil evidence might be explained by a relatively higher extinction rate in the Eastern Palearctic. Our integrative M4 model inferred an extinction event in this region following the Eocene-Oligocene cooling event (Fig. 2d), and during the Miocene, the extension of

climatically suitable area in the Eastern Palearctic was dramatically reduced (Fig. 2a), which may explain why this region is now scarcely represented among the major *Hypericum* lineages (Fig. 2d). Only a small portion on South East Asia, which now harbors the East Asian *Ascyreia* lineage (clade D), presented favorable conditions until the end of the Pliocene (Fig. 2a,d). Thus, *Ascyreia* could be a relict of a former Eastern Palearctic distribution rather than the result of a recent Miocene colonization ([23]). Fossils of *Hypericum* have been described from a mixed-mesophytic assemblage in Yun- nan, China ([22]), adding support to the role of East Asia as a climatic refuge for temperate plants during the Cenozoic ([11], [33]).

The interplay between climatic fluctuations and changing intercontinental geographic connections through the Cenozoic has determined the availability of "migration" routes to plants ([11]). Most biogeographic studies use information on past continental connectivity to model the dispersal probability between regions ([34]). However, our analyses suggests that "ecological connectivity" – the need for climatic conditions along the corridor to be within the climatic tolerance of the organism that disperses – is often more important than the presence of a physical bridge ([33], [12], [11]). Migration of temperate plants between Palearctic and Nearctic regions in the Tertiary involved one of two dispersal routes: the North Atlantic Land Bridge (NALB) or the Beringian Bridge ([3]). Our bioclimate models (Fig. 2a) showed that the NALB was never climatically favourable for *Hypericum*, and that most dispersal events in the Cenozoic must have involved Beringia. The Late Miocene cooling event interrupted this connection, but it was regained again in the warm Mid-Pliocene interval (Fig. 2a-V), when cool temperate deciduous forests expanded northward to the expense of the taiga and tundra vegetation ([27]). This is observed in our biogeographic reconstruction, with several dispersal and subsequent vicariance events between the Nearctic and Palearctic regions dated around the Pliocene (Fig. 2d). Also, our climate projections indicate that the southern temperate regions of Australia, Africa, and South America had favorable climatic conditions for *Hypericum* throughout its evolutionary history (Fig. 2a). This stands in contrast with the current low diversity of these regions, which, except for Africa, harbor few endemic species. Australia and South America became geographically connected to the Holarctic in the Late Tertiary, but the tropical equatorial belt probably acted as a climatic barrier, preventing *Hypericum* to colonize these regions (Fig. 2a). In contrast, the aridification process that affected Africa from the Miocene onwards and led to the replacement of tropical forests by woodlands and savannah ([35]) created an ecological corridor that allowed *Hypericum* lineages to migrate to the south (Fig. 2a,d). This connection was broken during the Pliocene, when an increase in aridification led to the formation of large deserts across North Africa and Arabia (Fig. 2a).

4. Conclusions

Integrating biogeographic and ecological evidence from extant and extinct taxa gives a more accurate reconstruction of *Hypericum* response to Cenozoic climate changes, showing that niche conservatism prevailed during the evolutionary history of the genus. *Hypericum* ancestors presented wide climatic tolerances, allowing them to cope with the rapid shifts in Cenozoic climate, and lineage divergence was achieved through specialization. Present diversity patterns were driven by the interplay between land connection and climatic adequacy along migration corridors. This integrative approach might be applicable to other widespread angiosperm lineages to explore the evolutionary response of plants to ancient climate change.

5. Materials and Methods

5.1 Distributional and climatic data.

Distribution data for extant Hypericum species (1033 records) were obtained from taxonomic monographs, herbarium collections, and online databases; fossil distribution data (143 records) were obtained from the

literature ([22]) and online resources (Table S3; see SI Text for more details). We grouped non-Pleistocene fossil records into two time slices: "greenhouse" lineages (Eocene, Oligocene, and Early Miocene, 68 records) and "coldhouse" lineages, spanning the Late Miocene and Pliocene (65 records). The boundary between greenhouse and coldhouse time slices was set at the Late Miocene cooling event (ca. 11.5 Ma; [1]). Climatic data for current conditions were extracted from WorldClim ([39]). For past scenarios, we used six Hadley-Centre GCM climate simulations that incorporate the effect of changes in atmospheric CO2 concentration; the latter has been shown to be a good proxy for variations in global temperatures ([36]). These simulations were generated with the same underlying algorithms and the same set of variables (monthly temperature and precipitation values at a resolution of 2.50 x 3.75 degrees), and represent major warming or cooling events in Earth history (see SI Text, [6], ([28]): a) Early Eocene Climatic Optimum (55 Ma, 1120 p.p.m CO2); b) the Terminal Eocene Event (35 Ma, 560 p.p.m); c) Early to Mid Miocene Climatic Optimum (400 ppm); d) Late Miocene cooling event (11 Ma, 280 p.p.m); e) Mid Pliocene warming event (3.6 Ma, 560 ppm); f) Preindustrial World (before 1900, 280 ppm).

5.2. Ecological Niche Modeling.

ENM results are highly dependent on the number of selected predictors ([37]). We first estimated the minimum set of climatic variables needed to explain Hypericum occurrences using an ecological-niche factor analysis (ENFA; [38]) with the terrestrial world as the background area (see SI Text). We then estimated maximum and minimum values for these climatic variables in all distribution points and transfer these conditions to the geographical space using a generalized intersection procedure to represent the potential distribution of *Hypericum*. This binary representation was transformed to a continuous one by calculating the scale-invariant Mahalanobis Distance (MD) from the average climatic conditions rep- resenting the lineage's hypothetical "climatic optimum" ([41]). This simple method based on geometrical set theory was preferred over more complex modeling techniques ([21]) because it requires only "presence" data and is thus more appropriate for the often spatially-biased fossil record. To assess the degree of niche stability versus niche evolution along the history of *Hypericum*, we estimated the climatic distance from the i) "extant" optimum based on current distributions, and ii) "green- house" and iii) "coldhouse" optima based on fossils, onto the six climatic simulations to build a continuous geographic representation model, using MD as a measure of climatic favourability. These models allowed us to identify areas in the underlying paleocontinent reconstruction with suitable climate for *Hypericum* during the relevant paleotime frame. In addition, we compared the geographic representations derived from greenhouse fossil data when projected onto the recent Late Miocene and Pliocene climate layers with the projection of the coldhouse climate optimum onto the Eocene and Early Miocene layers. To test the sensitivity of our analyses to the boundary selected between both time slices, especially in relation to the variable Miocene period ([1]), we repeated these analyses after removing all Miocene records. We used Kruskal-Wallis and multiple post-hoc comparisons to examine changes over time in the climatic conditions favourable to Hypericum for both the whole terres- trial Earth pixels and those pixels with fossil observations. Finally, we examined the variation in the climatic space defined by the two ENFA factors of each major clade in Meseguer et al ([23]) phylogeny in order to evaluate niche segregation within the genus, and tested for significance using non-parametric tests.

5.3 Biogeographic analysis.

Ancestral ranges and main migration events were reconstructed on the 114 specimen-three plastid dataset phylogeny of Meseguer et al. ([23]) using the parametric likelihood method Dispersal-Extinction-Cladogenesis ([42]). The world was divided into seven operational areas based on patterns of endemicity in *Hypericum* and the paleogeographic history of the areas. We used the maximum clade credibility tree from BEAST as input tree, but also run DEC over 1000 trees from the posterior distribution to account for uncertainty in topology. To examine the influence of fossil ranges and climatic (past and present) constraints in the inference of biogeographic

scenarios, we set up four analyses in order of increasing complexity: a) M1 used only distributions of extant species; b) M2 used present occurrences but incorporated also a paleogeographic model with four time slices (60-35 Ma, 35-10 Ma, 10-3.5Ma, 3.5-0 Ma), representing major tectonic and climatic events hypothesized to have affected the migration rate of temperate plants ([43], [34]); c) M3 incorporated the same paleogeographic model as well as information from present and fossil ranges. The latter was done by constraining the ancestral range of crown node *Hypericum*) to include the Eastern Palearctic region, where the Late Eocene fossil *H. antiquum* was found; d) M4 was identical to M3 but further incorporated the information from the modeling of ancestral climatic niches. First, the ancestral area of the crown node was constrained to include the Eastern Palearctic and Nearctic regions; the latter is not represented in the fossil record, but ENM projections predicted it to be part of the distribution of *Hypericum* in the Late Eocene. Second, we modified dispersal rates in the paleogeographic model to reflect variations in climatic adequacy across regions over time, such as the availability of transient migration corridors across unsuitable areas or climatic barriers that are specific to *Hypericum*. Thus, unlike the paleogeographic model in M2-M3, dispersal probabilities in M4 reflected both the physical and "ecological" connectivity between regions. This was interpreted as the existence of regions/pixels within the climatic tolerance of the group in a given time period (see SI Text for more details).

References

1. Zachos JC, Dickens GR, Zeebe RE (2008) An early Cenozoic perspective on green- house warming and carbon-cycle dynamics. Nature 451:279-283.
2. Wolfe JA (1975) Some aspects of plant geography of the Northern Hemisphere during the Late Cretaceous and Tertiary. Ann Mo Bot Gard 62:264-279.
3. Tiffney BH (1985b) The Eocene North Atlantic land bridge: its importance in Tertiary and modern phytogeography of the Northern Hemisphere. J Arnold Arbor 66:243-273.
4. Morley RJ (2003) Interplate dispersal routes for megathermal angiosperms. Perspec- tives Pl Ecol Evol Syst 6:5-20.
5. Willis KJ, MacDonald GM (2011). Long-term ecological records and their relevance to climate change predictions for a warmer world. Annu Rev Ecol Evol Syst 42:267-87.
6. Beerling DJ, Foxa A, Stevenson DS, Valdes PJ (2011) Enhanced chemistry-climate feedbacks in past greenhouse worlds. Proc Natl Acad Sci USA 108:9770- 9775.
7. Wen J (1999) Evolution of eastern asian and eastern North American disjunct distri- butions in flowering plants. Ann Rev Ecol Evol Syst 30:421-455.
8. Schneck R, Micheels A, Mosbrugger V (2012) Climate impact of high northern vegetation: Late Miocene and present. Int J Earth Sci (Geol Rundsch) 101:323-338.
9. Peterson AT, Soberón J, Sánchez-Cordero V (1999) Conservatism of ecological niches in evolutionary time. Science 285:1265-1267.
10. Crisp MD et al. (2009) Phylogenetic biome conservatism on a global scale. Nature 458:754-756.
11. Donoghue MJ (2008) A phylogenetic perspective on the distribution of plant diversity. Proc Natl Acad Sci USA 105:11549-11555.
12. Wiens JJ, Donoghue MJ (2004) Historical biogeography, ecology and species richness. Trends Ecol Evol 19:639-644.
13. Ronquist F, Sanmartín I (2011) Phylogenetic methods in historical biogeography. Ann Rev Ecol Evol Syst 42:441-464.
14. Stigall AL (2012) Using ecological niche modelling to evaluate niche stability in deep time. J Biogeogr 39:772-781.
15. Yesson C, Culham A (2006) Phyloclimatic modeling: combining phylogenetics and bioclimatic modeling. Syst Biol 55:785-802.
16. Schnitzler J, Graham CH, Dormann CF,Schiffers K, Linder HP (2012) Climatic niche evolution and species diversification in the Cape flora, South Africa. J Biogeogr 39:2201-2211.
17. Smith SA, Donoghue MJ (2010) Combining Historical Biogeography with Niche Mod- eling in the Caprifolium Clade of Lonicera (Caprifoliaceae, Dipsacales). Syst Biol 59:1-20.
18. Mao K, et al. (2012) The distribution of living Cupressaceae reflects the breakup of Pangea. Proc Natl Acad Sci USA 109:7793-7798.
19. Nauheimer L, Metzler D, Renner SS (2012) Global history of the ancient monocot family Araceae inferred with models accounting for past continental positions and previous ranges based on fossils. New Phytol 195:938-950.
20. Maguire KC, Stigall AL (2009) Using ecological niche modeling for quantitative bio- geographic analysis: a case study of Miocene and Pliocene Equinae in the Great Plains. Paleobiology 35:587-611.
21. Nogues-Bravo D, Rodriguez J, Hortal J, Batra P, Araujo MB (2008) Climate change, humans, and the extinction of the woolly mammoth. PLoS Biol 6:e79.
22. Meseguer AS, Sanmartín I (2012) Paleobiology of the genus Hypericum (Hyperi- caceae): A survey of the fossil record and its palaeogeographic implications. An Jar Bot Mad 69:97-106.
23. Meseguer AS, Aldasoro JJ, Sanmartín I (2013) Bayesian inference of phylogeny, mor- phology and range evolution reveals a complex evolutionary history in St John's wort (Hypericum). Mol Phylo Evol 67(2):379-403.
24. Nürk NM, Madriñán S, Carine MA, Chase MW, Blattner FR (2012) Molecular phylo- genetics and morphological evolution of St. John's wort (Hypericum; Hypericaceae). Mol Phylogenet Evol 66:1-16.
25. Davis CC, Webb CO, Wurdack KJ, Jaramillo CA, Donoghue MJ (2005) Explosive radiation of Malpighiales supports a mid-Cretaceous origin of modern tropical rain forests. Amer Nat 165:E36-E65.
26. Ruhfel BR et al. (2011) Phylogeny of the clusioid clade (Malpighiales): Evidence from the plastid and mitochondrial genomes. Am J Bot 98:306-325.
27. Salzmann U, Haywood AM, Lunt DJ, Valdes PJ, Hill DJ (2008) A new global biome reconstruction and data-model comparison for the Middle Pliocene. Global Ecol Bio- geogr 17:432-447.
28. Beerling DJ et al. (2012) Ecosystem CO2 starvation and terrestrialsilicate weathering: mechanisms and global-scale quantification during the late Miocene. J Ecol 100:31-41.
29. Peterson AT (2011) Ecological niche conservatism: a

timestructured review of evi- dence. J Biogeogr 28:817-27.

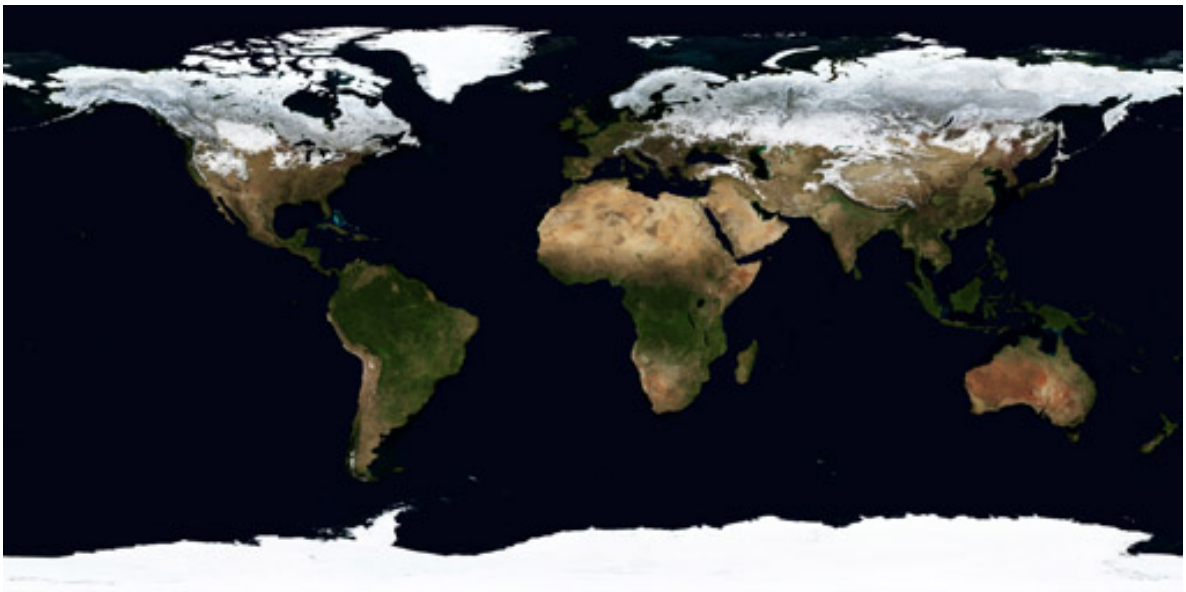
30. Linder HP, Hardy CR, Rutschmann F (2005) Taxon sampling effects in molecular clock dating: An example from the African Restionaceae. Mol Phylogenet Evol 35: 569-582.
31. Litsios G, Salamin N (2012) Effects of phylogenetic signal on ancestral state recon- struction. Syst Biol 61:533-538.
32. Lieberman BS (2005) Geobiology and paleobiogeography: tracking the coevolution of the Earth and its biota. Palaeogeogr Palaeoclimatol Palaeoecol 219:23-33.
33. Sanmartín I, Enghoff H, Ronquist F (2001) Patterns of animal dispersal, vicariance and diversification in the Holarctic. Biol J Linnean Soc 73:345-390.
34. Buerki S, et al. (2011) An evaluation of new parsimony-based versus parametric in- ference methods in biogeography: a case study using the globally distributed plant family Sapindaceae. J Biogeogr 38:531-550.
35. Senut B, Pickford M, S'ãgalen L (2009) Neogene desertification of Africa. Compt Rend Geosci 341:591-602.
36. Beerling DJ, Royer DL (2011) Convergent Cenozoic CO2 history. Nature 4:418-420
37. Beaumont LJ, Hughes L, Poulsen M (2005) Predicting species distributions: use of climatic parameters in BIOCLIM and its impact on predictions of species' current and future distributions. Ecological Modelling 186:250-269.
38. Hirzel AH, Hausser J, Chessel D, Perrin N (2002) Ecological-niche factors analysis: how to compute habitat-suitability maps without absence data. Ecology 83:2027-2036.
39. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. Int J Climatol 25:1965-1978.
40. Aarts G, Fieberg J, Matthiopoulus J (2011) Comparative interpretation of count, presence-absence and point methods for species distribution models. Methods Ecol Evol 3:177-187.
41. Varela S, Lobo JM, Hortal J (2011) Using species distribution models in paleobio- geography: A matter of data, predictors and concepts. . Palaeogeogr Palaeoclimatol Palaeoecol 310:451-463.
42. Ree RH, Smith SA (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst Biol 57:4-14.
43. Donoghue MJ, Smith SA (2004) Patterns in the assembly of temperate forests around the Northern Hemisphere. Phil Trans R Soc Lond B 359:1633-1644.

CHAPTER 5

The strategy of the marathon runner: ancestral resilience drove the evolution of the species-rich, age-old genus *Hypericum* (St John's wort, Hypericaceae)

This chapter has been done in collaboration with: Jorge M. Lobo (Museo Nacional de Ciencias Naturales, CSIC, Madrid), and Isabel Sanmartín (Department of Biodiversity and Conservation, Real Jardín Bótanico-CSIC, Spain)

The publication of this chapter is still on preparation.



CHAPTER 5

The strategy of the marathon runner: ancestral resilience drove the evolution of the species-rich, age-old genus *Hypericum* (St John's wort, Hypericaceae)

Andrea Sanchez Meseguer, Jorge M. Lobo, Isabel Sanmartín

A b s t r a c t

Hypericum originated in the Late Eocene, probably from tropical African ancestors, and dispersed to Eurasia and rapidly to the rest of the Holarctic at a time when tropical climates dominated northern latitudes. Initial diversification coincided with the dramatic cooling event at the end of the Eocene, implying a change of niche from tropical to temperate climates. Previous results have shown that the climatic niche of the genus has not changed for the last 35 Ma despite important temperature oscillations, although the different clades became specialized to a particular ecological space. Diversification rates have not been linked to ecological key innovation but rather reflect the steady accumulation of species through time and the climatic plasticity of *Hypericum* ancestors. Much interest has been focused on the study of adaptive radiations and key innovations to understand the generation of biodiversity. However, our results demonstrate that ancient lineages like *Hypericum*, which have survived through the climate changes of the Cenozoic, also merit attention for what they can tell us about the role of "resilience" and "plasticity" in coping with extinction.

1. Introduction

Over centuries, biologists have focused their attention on lineages that present extraordinary species richness (Schluter 2000; Linder 2008), trying to identify potential evolutionary novelties that mediated into their adaptation. The importance of novelties in evolutionary theory lies on the appeal of causally linking evolutionary change on individual traits with an increase in diversification rates (Heard & Hauser, 1995), since such correlation offers the possibility to investigate the causes of lineage accumulation and discover why some lineages diversify and others do not (Drummond et al., 2012). In flowering plants, there has been substantial research into the morphological, ecological and physiological correlates of cladogenesis (Hughes & Eastwood, 2006; Hodges, 1997; Waser, 1998; Sargent, 2004; Ree, 2005). Among them, ecological innovation has attracted special attention (Rosenzweig, 1995), in part for its double effect over cladogenetic (speciation and extinction) and biogeographic (dispersal) processes. Ecological niche differentiation has been associated with diversification (Graham et al., 2004; Moritz et al., 2000; Kozak & Wiens, 2007), and/or morphological change (Eldredge & Gould, 1972), being invoked as a causal explanation for adaptive radiations (Lovette, et al., 2002). Although the precise mechanism is still under debate, the evolution of ecological preferences may promote cladogenesis by influencing the dispersal capabilities of organisms (Wiens et al., 2010; Wiens & Donoghue, 2004), increasing or decreasing the geographic range of the species – widespread species might be more prone to speciate by allopatric speciation, and less likely to go extinct (Goldberg et al., 2011; McArthur & Wilson, 2011) – or by reducing competition through sympatric speciation (ecological displacement; Schniztler et al., 2012).

Ecological innovation ultimately determines the ability of the organisms to adapt to new environments. Under a scenario of climate change, organisms respond in two ways: conserving its ecological preferences (referred here as niche conservatism) or evolving ecological related traits to adapt to new environments. If evolutionary change occurs, then species could potentially persist in the original area. If not, they either go extinct or have to disperse to new areas that are within their climatic tolerance range (i.e., niche tracking). In many cases, this balance between niche conservatism and niche evolution is determined by the genetic plasticity of the organisms (Sgor et al., 2010), but also by the interplay between niche dynamics and the geographic template. On the one hand, niche preferences directly influence the ability of organisms to disperse and colonize new regions (Wiens et al., 2010). On the other, and under a scenario of niche conservatism ("it is easier to move than to adapt", Donoghue, 2008), species will disperse to more favourable areas if climatic corridors

exist (geographic sorting; Herrera 1992; Ackerly 2004). Conversely, if corridors are missing, then relevant adaptations to the new environment will presumably evolve (Donoghue, 2008).

In general, key innovation hypotheses are difficult to test because of the difficulty to establish causality (Cracraft, 1990), but also because evolution rarely affects a separate evolutionary trait in a single event. More often, the evolution of tolerances requires the adjustment of very complex physiological systems (Donoghue, 2008), with multiple linked changes occurring through time. Indeed, among all studies investigating evolutionary innovations, only relatively few have identified a correlation between shifts in diversification rates and the evolution of derived traits. Aside from a few well-explored model groups mainly from islands (Baldwin and Sanderson 1998), the mechanisms generating and maintaining diversity are poorly understood, and this gap is especially large for species-rich continental lineages (Losos, 2010) for which distributions are widespread and origins ancient. Indeed, many cases of key innovation driving diversification have been reported in young, recently diverged groups (Schniztler et al., 2012; Evans et al., 2009; Graham et al., 2004).

A comprehensive understanding of the causes of diversification will require to test key innovation scenarios against null hypothesis in which diversification is stochastically constant and any shifts in diversification are not significantly correlated with key innovations (Moore & Donoghue, 2007). The ideal group to test these hypotheses would be one that has experienced biotic or abiotic pressures through time, which presumably predisposed it to evolve derived adaptive traits, and which shows an unusually high species richness compared to its closest relatives.

One such group is the angiosperm genus *Hypericum* (Hypericeae). This genus, belonging to the "clusioid clade" of order Malpighiales (Rufhel et al., 2011) is considered one of 100 largest angiosperm genera in the world (Carine & Christenhusz, 2010). Most of this diversity – 357 out of its 496 species - is concentrated in the temperate Holartic region, although the genus extends also its distribution into the tropical and subtropical mountains of the Southern Hemisphere. Species of *Hypericum* occur in all continents, except Antarctica, and across a wide range of ecosystems, being absent from the poles, arid deserts, and low-altitude tropical areas. In addition, *Hypericum* stands out within the clusiod clade for its variety of life forms, ranging from trees to shrubs (the dominant form) and herbaceous plants. In contrast, the majority of clusioid genera are tropical woody plants inhabiting the rainforests of South America, Africa, and Asia, and they exhibit a more modest number of species (Table 1). Davis and collaborators postulated that the ancestors of the clusioid clade were Late Cretaceous tropical plants that inhabited close-canopy rainforests in Gondwana. Some of these lineages migrated into the Holarctic in the Early Tertiary when tropical conditions dominated the northern latitudes (Davis et al., 2005; Rufhel, 2011), as evidenced by the presence of clusioid fossils from the Late Cretaceous of North America (*Palaeoclusia*; Crepet & Nixon, 1998; Davis et al., 2005; Stevens, 2007), Early Tertiary of Siberia (*H. antiquum*; Meseguer & Sanmartin, 2012, Arbuzova, 2005), and Middle Eocene-Early Oligocene of Southern Europe (*Calophyllum* pollen in Spain; Cavagnetto & Anadón, 1995). During the course of the Cenozoic, these clusioid lineages would have gone extinct (i.e., extant species are now pantropically distributed), except for *Hypericum*, which survived and diversified in the temperate habitats of the Holarctic (Meseguer et al., 2013, submitted; Nurk et al., 2013). This makes *Hypericum* one of the few plant taxa that succeeded to make the difficult transition from tropical climates into the cool and highly seasonal environments of the northern latitudes (Judd et al., 1994; Wiens & Donoghue, 2004; Donoghue 2008).

Another aspect that makes *Hypericum* interesting is its old age. The last 65 Ma, the Cenozoic, have witnessed the origin and diversification of many major plant groups that exist today (Magallón & Sanderson, 2001). It was also a period characterized by major climatic changes. A greenhouse world with tropical climates at northern latitudes dominated the Early Cenozoic. This climaxed in the Paleocene-Eocene Thermal Event (PEET) at 55.8 Ma, one of the warmest intervals in Earth history. From this time onwards, the Cenozoic was characterized by a gradual decrease of global temperatures punctuated by some warm intervals (Zachos et al., 2008). The Terminal Eocene event (TEE, ca. 35 Ma) and the Late Miocene cooling (LMC, ca. 11 Ma) were probably the periods when more severe temperature drops occurred (Zachos et al., 2008), which have been associated with important changes on the vegetation cover of the Northern Hemisphere (Wolfe, 1975; Tiffney, 1985a,b; Morley, 2003; Willis & MacDonald, 2011). The TEE marked the decline of the Paleocene-Eocene "boreotropical" flora (Wolfe, 1975; Tiffney, 1985a,b), a uniform forest belt composed by evergreen and hardwood deciduous taxa that extended across the Northern Hemisphere in the Early Cenozoic through a narrower Atlantic Ocean and a warmer Beringia. Increases in cooling and aridity promoted the selection of the cold-temperate boreotropical elements and the expansion of a temperate deciduous vegetation, the "mixed-mesophytic forest", over the previous extension of the boreotropical flora (Tiffney, 1985ab). The LMC is associated with the replacement of the mixed-mesophytic vegetation by a boreal coniferous forest across the northern regions of Eurasia and North America (Schneck et al., 2012), while temperate plants became restricted to climatic refugia in the south. This is observed in the many

examples of temperate lineages with sister taxa disjunctly distributed in eastern North America and East Asia (Donoghue & Moore, 2003; Xiang et al., 1998, 2000; 2001; Wenn et al, 1999; Sanmartín et al, 2001). Wurdack & Davis (2009) suggested that all families of Malpighiales diverged in a rapid burst around the Mid-Late Cretaceous, and Meseguer et al (2013) dated the divergence between *Hypericum* and its closest relatives (stem-age) at 50 Ma and the first diversification event (crown-age) in the Late Eocene, c. 35 Ma. This implies that *Hypericum* or its stem lineages would have experienced all the major climate changes of the Cenozoic, including the PEET, TEE, and LCM events. Indeed, Meseguer et al. (2013) postulated that the initial diversification of *Hypericum* was promoted by a change of climatic tolerances to adapt to the much cooler environments after the TEE event (see also Nürk, 2011). Similarly, the transition from shrub to herbaceous forms that took place in several major clades around the Late Miocene has been considered an adaptation to the increasingly cooler and drier temperatures in the Northern Hemisphere, which could explain the rapid diversification and extraordinary species richness of the genus (Meseguer et al., 2013; Nurk et al. 2013).

Altogether, these characteristics, large species richness, ancient age, and presumed events of evolutionary change make *Hypericum* an ideal model system to study the interplay between niche dynamics, geographic evolution and species diversification in fuelling the generation of biodiversity. Two hypotheses can be considered. The "**key-innovation diversification**" (KID) hypothesis states that a change in climatic tolerances to deal with the new temperature regimes in the Holarctic and/or the appearance of the herbaceous habit form have been "adaptive breakthroughs", allowing the genus to diversify rapidly and being responsible for its evolutionary success. The null hypothesis, the "**time-to-speciation effect**" (TSE), suggests that rather than "rapid bursts of diversification", the old age of the genus and a steady accumulation of species through time might explain its large species richness. Predictions of the KID and TSE hypotheses imply different diversification correlates and trajectories, but they are not mutually exclusive on evolutionary change. Niche and morphological evolution could have occurred under both scenarios, but the difference lies on its particular effects over speciation and extinction rates: i.e., a significant change (KID) or no change (TSE).

Using fossil-based Ecological Niche Models (ENM), Meseguer et al. (submitted) inferred that climatic tolerances in *Hypericum* have remained stable for the last 35 Ma. They could, however, not test earlier periods (stem-*Hypericum*) due to the lack of relevant fossils, and limited their study to the pooled niche of the genus rather than focusing on major phylogenetic clades (Meseguer et al., 2013). Ancestral state reconstruction (ASR) techniques permit to infer ancestral ecological requirements without reference to the fossil record (Yesson & Culham 2006; Graham et al., 2004; Schnitzler et al., 2012; Smith & Beaulieu, 2009). At the same time recent developments in the field of macroevolutionary models allow now statistical testing of the association between geographic and niche evolution and diversification dynamics ("trait-dependent diversification models", Maddison et al., 2007; FitzJohn et al., 2009; Goldberg et al., 2011), while new episodic, time-variable diversification models permit to detect shifts in diversification rates correlated with major climatic changes (TreePar, Stadler 2011a,b).

Here, we used these stochastic birth-death models together with ancestral state reconstruction (ASR) of climatic niches, ENM projections, and a large-scale, dated phylogeny of *Hypericum* (and the clusioid clade) to understand the mechanisms that fuelled the engines of the genus extraordinary diversification. Unlike previous approaches (Graham et al, 2004; Evans et al., 2009; Vieites et al., 2009; Schniztler et al., 2012), we focus not on a single species or a small clade, but on a species-rich cosmopolitan clade, whose distribution spans nearly all continents and with an ancient origin extending back into the Early Cenozoic.. This allows us to study diversification and its drivers at different evolutionary scales, from the early history of the genus (as part of the clusioid clade) to the Late Cenozoic divergence of the major phylogenetic clades.

2. Material and methods

2.1 Phylogenetic and dating analyses

Bayesian relaxed molecular clock dating methods implemented in the software BEAST v. 1.6.1 (Drummond and Rambaut 2007) were used to infer phylogenetic relationships and absolute lineage divergence times. For *Hypericum*, we used the dated phylogeny published in Meseguer et al., (2013) and also employed in Meseguer et al. (submitted). This was based on DNA sequences of three chloroplast (*psbA-trnH*, *trnL-trnF* and *trnS-trnG*) and one nuclear (ITS) marker for 140 species, representing nearly 40% of the species diversity and including most of the described sections and morphological variation in the genus, as well as representatives of five Hypericaceae genera (*Eliea*, *Vismia*, *Harungana*, *Psorospermum* and *Triadenum*). For the clusioid clade, we used the dataset published by Ruhfel et al. (2011; available at TreeBase). This matrix contains different plastid (*matK*, *ndhF*, and *rbcL*) and mitochondrial (*matR*) markers and represents 71 of the 94 currently recognized clusioid genera (ca. 10% of the species diversity of the clade, including 21 *Hypericum* species). Divergence times within the clusioid clade were estimated using the uncorrelated lognormal relaxed clock model (UCLD) in a by-gene partitioned dataset using Yule tree prior and the GTR+G substitution model as additional

settings. Bayes Factor comparisons were used to select the appropriate model priors. Tracer v1.6 (Rambaut and Drummond, 2003–2009) was used to verify that all parameters have reached the stationary phase in log-likelihood values and the effective sampling size (ESS) for each parameter was > 200. A maximum clade credibility (MCC) tree with 95% high posterior density (HPD) intervals was generated with the program TreeAnnotator v. XX and visualized with Figtree v. 1.3.1. (Rambaut 2009). We included two calibration points used by Rufhel (2011) for estimating absolute ages: the Turonian (Late Cretaceous) fossil *Palaeoclusia chevalieri* (Crepet and Nixon, 1998) was used to constrain the age of the root node (i.e., the crown age of the clusioid clade) following Davis et al. (2005), using a lognormal calibration prior with the following parameters: offset or minimum hard bound = 89.3 Ma, mean = 0 and standard deviation (Std) = 0.9, with the 95% of the prior weight falling within the geological interval of the fossil (89.3–93.5 Ma). The Eocene fossil *Pachydermites diederexii* was used to constrain the age of crown Symphonieae, with a lognormal prior (40.4 – 48.6Ma), offset = 40.4, mean = 4.5, and Std = 0.3. Circumscription and placement of these fossils is discussed in Ruhfel (2011). In addition, *Hypericum antiquum*, from the Late Eocene of West Siberia (Arbuzova, 2005; Meseguer & Sanmartín, 2012) was used to constrain the crown node of genus *Hypericum*, with a lognormal prior (33.9-37.2 Ma), offset = 33.9, mean = 0 and Std = 0.7). Circumscription and placement of this fossil is discussed in Meseguer et al. (2013). We adopted the uppermost limit of the relevant time interval as the minimum hard bound and follow Walker & Geissman (2009) for the age of the geological periods.

2.2 Diversification tests

We used the above dated phylogenies and a diverse array of diversification methods to estimate changes in diversification rates in *Hypericum* and within the clusioid clade. We used the **gamma statistics** (Pybus & Harvey, 2000), corrected for incomplete sampling with the Monte Carlo constant rates test (MCCR), to test whether the pattern of lineage diversification departs from the constant rate birth-death model. To detect temporal shifts in diversification rates in the *Hypericum* phylogeny we used novel birth-death likelihood methods implemented in the R package **TreePar** (Stadler, 2011a,b, 2012). In particular, we make use of episodic birth-death models in which diversification rates are allowed to change at specific points in time (rate-shifts), while conditioning on the number of extant taxa and incorporating the effect of random incomplete taxon sampling. Maximum Likelihood (ML) optimization was used to simultaneously estimate diversification parameters - the net diversification rate ($r = b - d$, being b = birth and d = death) and the extinction fraction ($\varepsilon = d/b$) – for each time interval together with the rate-shift times (Stadler, 2011). TreePar cannot estimate simultaneously the shift times for all time intervals but instead uses a greedy algorithm that fixes one ML time shift before estimating the next one, and so on. We used likelihood ratio tests to compare nested models of increasing complexity with 1 to 8 additional rate shifts (an arbitrary high value based on the size of our phylogeny). The *Hypericum* phylogeny was divided into a grid of 0.33 Ma time intervals, steps between 1 and 30 Ma, as specific point times where turnover and diversification rates can be estimated (Standler, 2011). Additionally, TreePar can incorporate the effect of incomplete taxon sampling at any time interval. The sampling fraction for the first time interval (present) was set to 0.22 according to the taxon sampling in the present, while the analysis was repeated under two extreme values of survival probability for the past time intervals (this is interpreted as the effect of punctual (mass) extinction events removing some lineages from one interval to the next):

1) We assume that no extinction occurred in any period, 100% of the species survive from one period to the other (the sampling of deep branches is complete).

2) We assume that very high extinction occurred in the past and only 10% of the species survived between periods.

Time-variable diversification methods such as TreePar do not account for changes in diversification rates across clades, or for the uneven taxon sampling within different clades. Rate heterogeneity across clades may reflect the existence of trait-dependent speciation and extinction rates. To test this, we used the phylogenetic-taxonomic method implemented in the R package **turboMEDUSA** (Brown et al., 2012), which is based on the MEDUSA approach developed by Alfaro et al. (2009), to test whether there have been shifts in diversification associated to particular clades in *Hypericum* or within the clusioid clade. This method applies a stepwise AIC to identify rate shifts in diversification rates across time and clades. Incomplete sampling was incorporated using a taxonomic approach where tips of the phylogeny represent taxonomic groups with assigned species richness. Living diversity for every *Hypericum* clade was assessed based on phylogenetic results (Meseguer et al., 2013; Nürk et al., 2013) and the sectional classification of Robson (1977, 2012). For clusioid genera, species diversity was compiled based on the phylogenetic results of Ruhfel et al, (2011) and taxonomic revisions (Stevens, 2007). Extinction rates are notoriously difficult to estimate from phylogenies containing only extant taxa (Rabosky, 2009; Quental & Marshall, 2010). Therefore, we also used the **method-of-moments estimator** of Magallón & Sanderson (2001) implemented in the R package *geiger* (Harmon et al., 2008), to estimate the net diversification rate for a given clade under varying extinction fractions. This method does

not require a dated phylogeny but compares the age and size of a clade to estimate the net diversification rate under a constant birth-death process. We used two extreme extinction fractions: ε = 0 (no extinction) and ε = 0.99 (high extinction) following Magallón & Sanderson 2001 and Linder (2008).

2.3 Trait-dependent diversification models

Correlations between shifts in diversification rates and climatic or cladogenetic events do not imply causality. To explore if evolutionary change directly influences diversification dynamics, we used trait-dependent diversification models implemented in the R package *diversitree* (FitzJohn, 2010). In these models, the trait itself (transitions between character states) can influence the birth-death process that generates speciation times in the phylogeny (Maddison et al., 2007). We used these models to explicitly address the relationship between trait evolution (e.g., niches, geographic distribution, habit form) and diversification. First, we used the Binary State Speciation Extinction model (BISSE; Maddison et al., 2007; FitzJohn et al., 2009) implemented in *diversitree* to test whether a change of habit from woody to herbaceous forms can explain differences in extinction and speciation rates among major *Hypericum* clades. To account for incomplete taxon sampling, we used the pruned phylogeny from the turboMEDUSA analysis and applied an unresolved clade information approach (diversitree manual), which is equivalent to the taxonomic approach of turboMEDUSA, where tips of the phylogeny represent taxonomic groups with assigned species richness including the proportion of species for each character state (Fitz-John et al. 2009). Coding of traits was based on Robson's taxonomic monograph (Robson 1977 onwards), assigning both tree and shrubby habits to the "woody" trait.

Second, we used the Geographic State Speciation and Extinction model (GeoSSE) (Goldbert et al., 2011) to answer the question if living in certain geographic regions increases or decreases the rate of speciation and/or extinction. GEOSSE extends the BISSE model to incorporate a third, polymorphic state for geographic characters, since taxa are often not endemic but present in more than one area/state. This method accounts for random incomplete taxon sampling within areas and associates constant-rates birth-death models with a three-state Markov model. Estimated parameters are the within-region speciation (*s*A, *s*B), between- regions speciation (*s*AB), within-region extinction (*x*A, *x*B) and dispersal (*d*A, *d*B). Biogeographic areas were defined according to Meseguer et al. (2013), except for two areas that were excluded from the analysis. The Neotropics ("NT") only includes endemic species, and therefore cannot be analysed with GEOSSE, while the occupancy of Oceania ("OC") by *Hypericum* species is marginal. Species distributions were obtained from Robson (1981–onwards). For the selected geographic regions we estimated ML parameters for a full GEOSSE model (in which speciation, extinction and dispersal rates parameters differ between areas), but also for a set of GEOSSE constrained models: constrained within region speciation (*s*A ~ *s*B, *s*AB ~ 0), between region speciation (*s*AB ~ 0), dispersal (*d*A ~ *d*B) and extinction (*x*A ~ *x*B). Model fit comparison was assessed with likelihood ratio rest (R package *diversitree*) and comparing AIC values of the different models. In all the cases the full model better fitted the data than the constrained models. However, in some cases, such differences between models were not significant. The eight parameters estimated in the full GEOSSE model under ML were used as a prior for a Bayesian MCMC search. The MCMC chain was run for 10.000 generations, and the first 1000 were discarded.

Finally, we used GEOSSE to investigate differences in diversification rates between clusiod genera or *Hypericum* species living in tropical or temperate climates. We consider tropical as the taxa distributed between the tropic of Cancer and the tropic of Capricornio (23º of latitude North and South). Incomplete taxon sampling was accounted for using a similar approach as in BiSSE. Major biome classification was based on Stevens (2007) for the clusioid-clade and Robson (1977-onwards) for *Hypericum*. As in the previous case, we estimated ML parameters for a full GEOSSE model, but also for a set of GEOSSE constrained models, and assess model fit comparison with likelihood ratio rest and AIC values.

2.4 Ancestral niche reconstruction and geographic projections

We used taxonomic monographs (Robson 1981, 1985, 2012), herbarium collections, and online databases (GBIF), the latter collated to exclude ambiguous citations, to construct a database of 1033 occurrences (901 localities) (Meseguer et al., 2014). Climatic variables important to explain *Hypericum* occurrences were discriminated using an ecological-niche factor analysis (ENFA; Hirzel et al., 2002; Calenge & Basille, 2008) with the terrestrial world as the background area and 19 climatic variables for current climatic conditions (≈1950-2000) extracted from WorldClim (Hijmans et al., 2005). Further details can be found in Meseguer et al. (submitted). This procedure derived seven climatic variables representing important vegetation predictors: annual precipitation, annual variation in precipitation, aridity, continentality, maximum monthly precipitation, maximum monthly temperature, mean annual temperature, minimum monthly precipitation and minimum monthly temperature. Bioclimatic indices as aridity and continentality were calculated following the formulas provided by Valencia-Barrera et al. (2002). Maximum and minimum values for the selected variables were estimated for species grouped according to the phylogenetic clades recovered in Meseguer et al. (2013).

To reconstruct ancestral climatic preferences for major phylogenetic lineages within *Hypericum*, maximum and minimum values and tolerance ranges (max-min values) were estimated for the selected variables for each of the five phylogenetic clades (A to E) discussed in Meseguer et al. (2013). We used the function *getAncStates* in the R package *geiger* (Harmon et al., 2008) to optimize these variables onto Meseguer et al. (2013)'s phylogenetic tree, pruned to include only one tip per clade and outgroup taxa, using ML algorithms under a Brownian motion model (Schluter et al., 1997). To assess the magnitude of niche differentiation among clades, the scale-invariant Mahalanobis Distance (MD) was calculated from the average inferred climatic conditions (representing the lineage's hypothetical climatic optimum; Varela et al., 2011) of the ancestor of all Hypericaceae genera (root of the tree, node 14) with respect to each major clade. To better visualize the trajectory of niche evolution across clades, we also plotted the average inferred from the optimized maximum and minimum values of the least correlated variables, minimum monthly precipitation and the minimum monthly temperature.

Finally, to represent the potential distribution of the ancestral nodes in the past, we used a generalized intersection procedure by transferring the inferred maximum and minimum values for the selected climatic variables of each major phylogenetic clade to the geographical space in the past. For this, we used the same six paleoclimate simulations used by Meseguer et al. (submitted), which incorporate the effect of changes in atmospheric CO_2 concentration and represent major climate changes in the history of Earth (Beerling et al. 2009; Beerling et al 2011; Beerling et al. 2012 and Bradshaw et al. 2012): Early Eocene (PEET, 55.8 Ma); Late Eocene (TEE, 35 Ma); Early-Mid Miocene (before the Mid Miocene Climatic Optimum, 15 Ma); Late Miocene climate cooling event (LCM, 11 Ma); Mid Pliocene warming event (MPW, 3.6 Ma); Preindustrial World (before the industrial revolution ≈1900).

3. Results

3.1 Diversification analyses

Figure 1 shows the dated phylogeny (MCC tree) of the clusioid clade derived in BEAST. The topology of the tree is overall congruent with the one presented in Rufhel et al. (2011), in recovering all major families, but there are important differences in the inferred relationships among them. We have recovered family Podostemaceae sister to the other clusioid families; where Calophyllaceae is sister to Clusiaceae-Bonnetiaceae in our study, while Ruhfel et al. (2011) recovered Calophyllaceae sister to Hypericaceae-Podostemaceae. Attempts to constrain the topology to follow Rufhel et al. (2011)'s tree did not change these results. The cumulative number of lineages over time shown in the LTT plot suggests a constant birth-death process. A similar result was found for the dated *Hypericum* phylogeny (Figure 2). The gamma test, which is based in a pure birth process, does not reject a constant diversification model, even when the statistic is corrected for incomplete taxon sampling (-1.08, p =0.66). TreePar (Fig. 2) identified a significant rate shift at 9.25 Ma under the two extreme survival models: no extinction (LH=256.5, p=0.038) and high extinction (LH=256.5, p=0.038; Table 2). An additional rate-shift at 27 Ma was inferred under the model with no extinction and at 7.27 Ma under the high-extinction model, but none of these models were significant when compared with the corresponding one rate-shift model (LH=254.4, p=0.25, no extinction /LH= 255.13, p=0.43, high extinction). We also obtained nearly identical parameter values under these two extreme survival probabilities (r_1=0.197; ε_1=0.49; r_2= 0.045; ε_2= 0.93; Table 2).

The MEDUSA analysis selected a four rate-shift model (LH=-367.8, AICc = 768.5) for the clusioid clade (Fig. 3). The background tempo of diversification for the clusioid clade has low net rate (r=0.071) and moderate turnover (ε =0.593), and shift rates were identified in the ancestors of *Clusia* (r=0.476, ε =0.605), *Calophyllum* (r=0.432, ε =0.605), *Symphonieae* (r=0.0871, ε =0.0522), and the MRCA of Calophyllaceae-Clusiaceae (r=0.051, ε =0.945; Fig. 3). MEDUSA selected a two rate-shift model (LH=-85.99, AICc 195.98) for the *Hypericum* phylogeny (Fig. 4), with a significant change in diversification rates compared to the background diversification rate in the genus (r=0.059; ε =0.959) in two lineages: *Elodes*/*Adenotrias* (r=0.035, ε =2.65 x 10^{-07}) and the *Androsaemum* group (r=0.092, ε =6.02 x 10^{-8}). Collapsing the tree into unresolved branches depends to a certain extend of arbitrary criteria and can introduce error in the inference. We alternatively repeated the analysis under different topologies ranging from more to less conservative assignation of species richness to unresolved parts of the tree both in the clusioid and in the *Hypericum* matrices; we also repeated the analysis with and without outgroup taxa. All the analyses produced equivalent results.

Table 1 shows the absolute rate of diversification for major clusioid and *Hypericum* lineages, estimated with the method-of-moments estimator in the absence of extinction (ε = 0) and under a high relative extinction rate (ε = 0.99). Within *Hypericum*, the lineages with the highest net diversification rates were *Hypericum* s.l. *Brathys*, and *Ascyreia*, while *Webbia*, *Campylopus* and *Elodes* present extremely low diversification rates. Within the clusioid clade, *Hypericum* presents a slightly higher diversification rate than the whole clade, but other genera like *Calophyllum* or *Clusia* have experienced much bigger diversifications given their age (Table 1). However, these

are all young genera. When compared with other clusioid lineages of similar age (60-40 Ma), such as *Terniopsis* (6 species) or *Weddellina* (1 species), the diversification rate of *Hypericum* is actually high (Table 1).

3.2 Trait-dependent diversification models

BiSSE found not significant differences between the full model, in which all parameters differ between character states, and the unconstrained models ($p > 0.05$), suggesting that that there are not changes in speciation and extinction rates associated with being woody or herbaceous in *Hypericum*. GEOSSE, however, found that for all geographic regions the full model fitted the data better than the constrained models, suggesting the diversification in *Hypericum* has been range dependent, and speciation and extinction rates vary across areas. However, in some cases, such differences between models were not significant (Supplementary Material Table S1). Parameter values estimated for the different geographic regions are not directly comparable between areas, since every analysis was run independently (evaluating a single region against the rest of the world). Nevertheless, our MCMC Bayesian results (Fig. 5) indicate that *Hypericum* lineages from the Eastern Palearctic region (EP) have significantly higher extinction and speciation rates (species turnover) than the other geographic areas pooled together (Table S1, Fig. 5). This pattern has not been found for the other regions. Only Africa (AF) shows posterior extinction rates slightly higher than in the rest of the world, but mean values are very close and there is considerable overlap (Fig. 5). In the Nearctic (NE), Western Palearctic (WP) and Irano-Turanian-Himalayan (ITH) regions, speciation and extinction rates are lower than in the rest of the world with different degrees of overlap (Table S1, Fig. 5). Dispersal rates from each region to the rest of the world (*d*A) were higher than in the opposite direction (*d*B) in all cases (Table S1). The between-regions speciation parameter (*s*AB) is high in all cases, suggesting a high frequency of vicariance and peripatric isolation events in the history of the group (Golberg et al, 2011). The highest value is found in the ITH region, which also presents a high dispersal rate *d*A.

For the analysis of diversification rates between clusioid genera living in living in temperate and tropical regions, GeoSSE favoured a constrained model in which speciation is constrained to be equal among areas over the full model with different speciation and extinction rates (LH=-731.18, AIC=1472.4). Under this model, extinction rates associated with tropical taxa (A) are slightly higher than those associated to temperate (B) groups (*s*A 0.773, *x*A 0.724, *x*B 0.689). For *Hypericum*, GEOSSE selected a full model (LH=-421.92, AIC=857.84), showing higher speciation and extinction in tropical regions (*s*A 0.767, *s*B 0.233, *s*AB 0.730, *x*A 0.743, *x*B 0.078, *d*A 0.0, *d*B 0.122).

3.3 Ancestral niche reconstructions

Table 3 summarizes maximum and minimum values for the selected variables optimized in *geiger* for each major node in Fig. 4; Table 4 shows the same optimizations for the tolerance range. Present values for these variables can be found in Tables S2 and S3. The ancestral niche reconstruction along the phylogenetic history is shown in Figure 6. The most striking result is the large MD distance between the inferred climate optima for the ancestor of Hypericaceae (node 14, Fig. 6a) and the next node, the ancestor of tribes Vismieae-Hypericeae (stem-node *Hypericum*: nodes 14 and 15 respectively, Fig. 6a), to the climate optimum of the most recent common ancestor (MRCA) of all living *Hypericum* (crown-node 18). Interestingly, a similarly high MD distance (a measure of dissimilarity) was found between these two ancestral nodes (14, 15) and the MRCA of Vismieae (node 16). A slight increase in climatic tolerances (the size of the circle) can also be observed over time (Fig. 6a).

The MD, however, only indicates the magnitude of the difference (distance) between niche values, but does not provide information about which clades are more similar or dissimilar to each other. To investigate this, we plotted the optimized ancestral values of the minimum monthly temperature against minimum monthly precipitation for every node (Table 3, Fig. 6b,c) – these variables present the lowest correlation value (Table S3) and are often used to characterize climatic regions of the world (Kottek et al., 2006; McKnight & Hess 2000). Again, the most striking result is that although there is a general trend towards aridity (lower minimum precipitation) and decreasing minimum temperatures along the evolution of *Hypericum*, there are also numerous departures and approaches of the optima for the crown nodes of present *Hypericum* lineages from and towards the ancestral conditions represented by the basal nodes 14-15 (stem-*Hypericum*) and node 18 (crown-*Hypericum*). Analyzing the values in more detail, the hypothetical climatic optima of ancestors of Hypericaceae and Vismieae-Hypericaceae (nodes 14 and 15 respectively) presented an average minimum temperature of 10ºC in the coldest month and 65–70 mm of precipitation in the driest month. The ancestors of Vismieae (nodes 16, 17) present a considerably higher estimate of the minimum monthly temperature values, above 15ºC, but seem to have lived under lower precipitation regimes (a minimum of 55–60 mm in the driest month, Fig. 6b-c). The group clustering the majority of lineages ascribed to genus *Hypericum* (nodes 18 to 25) presents similar precipitation values than the ancestors of Vismieae, but they could live under much lower temperature regimes, ranging from 5ºC to 0ºC in the coldest month. The ancestor of the *Ascyreia* s.l. clade (node 24) is the main exception, as it seems to have thrived

in much more humid environments, over 75 mm of precipitation in the driest month (Tables 3-4, Fig 6b-c). Plotting the present optima for the major phylogenetic clades (the "red squares" in Fig. 6c) shows a similar pattern to the one observed in the ancestors, with several approaches and departures relative to the ancestral niche represented by crown node 18. The early-diverging clade *Elodes-Adenotrias* and the *Androsaemum*-group show average values very similar to those of crown group *Hypericum*. Again, the opposite pattern is represented by *Ascyreia* –and its sister-group *Campylosporus* – which show the most divergent values with respect to the ancestral conditions. The latter also represents a case of "ecological vicariance", in which sister-taxa evolve in opposite directions along two climate dimensions: *Campylosporus* in warmer and drier conditions, *Ascyreia* in more humid but colder environments. A similar example is the sister-pair *Brathys-Myriandra* (Fig. 6c).

The projection of the inferred ancestral niche values onto the paleoclimate-geographic scenarios (Fig 7) shows an overall continuous occupancy of the southern parts of the Northern Hemisphere (the eastern and western coasts of the Tethys Sea), as well as the southernmost regions of the Southern Hemisphere. The extension of the potential area increases from the past to the present, with the largest area shown by the ancestors of the Old World group (nodes 22–25). The ancestors of tribe Vismieae show a different pattern, with a predicted distribution in tropical areas of Central and South America, Africa and South East Asia-Oceania (Fig. 7).

4. Discussion

4.1 Early diversification history of Hypericum

Hypericum is one of the few plant taxa hypothesized to have been able to evolve tolerances from tropical ancestors to the cooler and highly seasonal habitats of temperate latitudes (Donoghue, 2008; Meseguer et al., 2013; Nurk et al., 2013). Yet, previous analyses have failed to identify if or when this transition took place. Meseguer et al. (submitted) found that climatic tolerances in *Hypericum* have remained relatively constant since the crown group diversification, 35 million years ago. They based their inference on a fossil-informed approach to ecological niche modelling and biogeographic analysis, so they could not go further back than the earliest fossil record of the genus, *Hypericum antiquum* from the Late Eocene (Meseguer & Sanmartín 2012; Meseguer et al., 2013). Davis et al. (2005) used ASR techniques to infer that the ancestors of family Hypericaceae inhabited open areas of the tropical forest, probably occupying open or degraded parts of the boreotropical belt, which suggests that *Hypericum* ancestors could have some form of preadaptation to temperate conditions before their entrance into the Holarctic and the drastic drop of temperatures of the TEE. However, Davis et al. (2005) inference was limited by the low taxon sampling (only Hypericaceae genera *Hypericum* and *Vismia* were included in the reconstruction). Results from our ASR analysis suggest that there has actually been a change of niche between *Hypericum* stem lineages – the ancestor of Vismieae-Hypericeae (node 15) - living under relatively warm and humid conditions (average minimum monthly temperature of 10ºC and precipitation of more than 60 mm), and crown group *Hypericum* (node 18), which inhabited colder and slightly drier environments (Fig 1a). More interesting is the fact that the ancestors of Hypericaceae and stem-*Hypericum* (nodes 14-15) show climatic preferences that are intermediate between those inferred for temperate *Hypericum* (crown node 18) and the more tropical Vismieae (nodes 16-17; Fig. 6a-c; Table 2-3). This could support Davis et al. (2005)'s inference and suggest that ancestors of Hypericaceae (and stem-*Hypericum*) probably lived under lower temperatures and higher precipitation regimes than present tropical clades. One explanation for this is that this suit of climatic values reflects the preferences of a boreotropical group, as suggested by Meseguer & Sanmartín (2012). During the warmth peak of the Paleocene–Eocene, temperatures in the Holarctic have been described to be between 5 and 10 degrees Celsius warmer than present temperate conditions (Zachos et al., 2008), which is in agreement with the average minimum temperature values inferred for the ancestors of Hypericaceae (Fig. 6b-c). The boreotropical forest of the Early Cenozoic was composed of a mixture of evergreen subtropical elements and hardwood deciduous taxa, which no analogy to any of the forests we see today (Wolfe, 1969, 1972, 1975; Tiffney 1985ab; Sanmartín et al., 2001), which suggest that the region presented particular climatic features not found nowadays in tropical or temperate latitudes. One msethodological flaw of the Brownian model used in our ASR reconstruction is the tendency to infer intermediate characters for the ancestors as we move to the past. An intermediate estimate might reflect that the ancestor is a generalist, but it may also reflect uncertainty - the ML estimator of an ancestor for a continuous trait is simply the weighted average of the dimensions of the extant species at the tips of the tree (Schluter et al. 1997) – as well as the effect of extinction removing early-diverging branches, and associated phylogenetic information, as we get closer to the root node. Nevertheless, the fact that our niche-geographic projections for the ancestral nodes 15 and 18 coincide with those inferred by Meseguer et al. (submitted) based on independent evidence from the fossil record lends, support to our ASR reconstructions. Moreover, the idea of a generalist, intermediate ancestor agrees well with Meseguer et al.'s finding that early lineages in *Hypericum* had broad climatic tolerances, which allow them to live in a wide range of climatic worlds.

The onset of the *Hypericum* crown radiation coincided with the drastic temperature decline at the end of the Eocene (TEE; Meseguer et al., 2013, Fig. 2). This has been associated with a change in niche preferences followed by a rapid diversification (Meseguer et al., 2013; Nurk et al., 2013). Although our results suggest a change of climatic niche, there does not seem to have been a concurrent increase in diversification rates. The pattern of lineage accumulation in *Hypericum* (LTT plot, Fig. 2) does not fit what is expected from an event of rapid diversification or adaptive radiation. As ecological niches are being filled, the initial rapid diversification slows down, observed as a significant decrease in diversification rates over time ("density-dependent-cladogenesis", Rabosky & Lovette 2008). This is not the pattern exhibited by *Hypericum*, according to the gamma test. If any, there has been an increase in diversification rates over time (Fig. 2), with TreePar recovering a significant acceleration at 9.25 Ma. Moreover, the current, extant diversity of *Hypericum* is not particularly high, given its age and the background diversification rate of the clusioid clade as shown by the method-of-moments estimator (Fig. 3, Table 1). This stands in contrast with younger tropical clusioid genera like *Clusia* or *Garcinia*, which have experienced remarkable radiations. Although *Hypericum* is the clusioid genus with more species, it is also one of the oldest (Table 1). This suggests that steady accumulation of lineages during the last 35 Ma might explain better the species richness of the group than an initial rapid radiation. Nonetheless, *Hypericum* is richer than other clusioid clades of the same age (Table 1), and we found that extinction rates were higher for tropical clusioid genera than for those living in temperate latitudes. This might indicate that the evolution of temperate tolerances in *Hypericum* was not so much associated to an increase in speciation rates but a decrease in its extinction risk. Besides evolution of tolerances to colder environments, the apparition of the herbaceous form has been associated to the current large number of species in *Hypericum* (Meseguer et al., 2013; Nürk, 2011). Although this transition has been argued to increase the rate of niche and geographic evolution of clades (Smith & Bealieu, 2009), we did not identify differences in diversification rates associated with being woody or herbaceous in *Hypericum*. In sum, all evidences seem to indicate that the appearance of evolutionary innovations, such as new tolerance to temperate climates or the appearance of the herbaceous form, did not increase the rate of cladogenesis in *Hypericum*; if any they might have contributed to decrease its extinction risk. Some authors have argued that the tempo of diversification can be inherited among lineages, probably by the heritability of lineage-specific traits that directly affect diversification rates, such as the geographic ambit or the extinction risk (Savolaine, 2002; Rabosky, 2009).

Another interesting result from our ASR reconstructions is the inference that *Hypericum* stem lineages (nodes 14-15) inhabited geographic areas that were nearly identical than those of their descendants, suggesting very little geographic evolution despite a change in climate niche. Geographic projections in Fig. 6 indicate a consistent coverage of the margin of the Thethys Sea, the southern parts of North America and a small portion of South East Asia in the Northern Hemisphere, and the southernmost Southern Hemisphere from the beginning of the Cenozoic until the present. In fact, the geographic projection for node 18 is overall comparable with the fossil-based ENM reconstructions of the Late Eocene published by Meseguer et al. submitted: SI Fig 3 a,g), although the predicted distribution is smaller in our study, not including the Asian portion of Beringia, the central regions of the EP and the northernmost areas of Europe (Fig 1d; Meseguer et al, submitted SI Fig 3). The reduction in the potential past distribution when comparing projections based on ASR or fossil-based ENM models could be explained by the effect of extinction: because ASR is only based on present occurrences, the tolerances of extinct taxa are not considered in the analysis and therefore, the amount of information is reduced as we move deep into the past. Although direct comparisons of more recent ancestral reconstructions is not straightforward, since Meseguer et al. (submitted)'s fossil niche projections only considered the pooled niche of the genus, it is remarkable that our ASR projections shows *Hypericum* ancestors living in areas where *Hypericum* fossils actually occur. Specifically, the potential area in node 18 (*Hypericum* crown group) includes an eastern portion of the Turgait strait, where the oldest fossil remains of *Hypericum*, *H. antiquum* and *H. septestum*, were found (Meseguer & Sanmartín 2012). Likewise, the Early Miocene predictions for nodes 23-24 included the area of appearance of *Hypericum* fossil remains in Yunnan, in a mixed-mesophytic forest assemblage (Zhao et al., 2004). In contrast, the predicted area of occurrence for the ancestors of the New World group and *Ascyreia* is larger than the actual distribution of these clades (Fig. 7), suggesting that other factors in addition to climate have determined the geographic distribution of these groups. Another example is the southernmost distribution of *Hypericum* in the temperate regions of the Southern Hemisphere. Even if the area was climatically suitable throughout the history of the genus, there are no extant *Hypericum* species distributed in these regions, probably because the tropical intermediate regions in these continents (Neotropical South America, South East Asia) acted as a climatic barrier. The only exception is South Africa, where the formation of an aridity corridor in the Late Miocene allowed some clades to move southwards (Meseguer et al. submitted).

4.2 Evolutionary dynamics within Hypericum

Niche dynamics can vary across temporal and phylogenetic scales (Wiens et al., 2010). At the genus level, *Hypericum* tolerances evolved at the end of the Eocene and have remained relatively stable from this time onwards (Fig. 6-7; Meseguer et al., submitted). This was accompanied by geographic stability, with *Hypericum* ancestors inhabiting to a great extent the same regions than their extant descendants (Meseguer et al., submitted, Fig. 7). On the other hand, Meseguer et al. (submitted) reported significant niche segregation among current major phylogenetic clades along the composite "aridity" and "tropicality" factors: the more "tropical" New World and Asian *Ascyreia* groups (clades B and D, Fig. 4) were segregated from the more "xeric" European and African clades A (*Elodes-Adenotrias*), C (*Androsaemum*) and E (*Hypericum*). They explained this apparent contradiction as the result of the different phylogenetic clades specializing in different subsets of the broad generic niche rather than as change to new environmental conditions (niche evolution). Our results based on the phylogenetic-based reconstruction of ancestral niches lend some support to this assertion. As in Meseguer et al.'s study, phylogenetic niche variation seems to have occurred mainly along the temperature gradient, while precipitation has been less labile for the majority of clades (Fig. 6b). There is a clear, observable tendency towards increased cold tolerance along the evolution of *Hypericum*, from the Early Eocene boreotropical ancestors (stem-*Hypericum*, node 15), living under an average minimum monthly temperature of 10ºC, to the cold-temperate ("mixed-mesophytic) crown-group ancestor (node 18) at 4ºC, to the present clades, such as *Hypericum* s.l or *Hirtella* s.l, which present minimum monthly temperatures under 0ºC (Fig. 6b-c). A weaker trend towards increasing drought resistance can also be observed along the lineages of *Hypericum* and Vismieae with respect to their ancestors (nodes 14-15, Fig. 6b). This agrees well with the global cooling trend that started by the Mid Cenozoic, promoting the appearance of cooler and more seasonal environments in the Holarctic (Zachos et al., 2008) and a trend towards increased aridification in Africa (Plana 2004).

However, even more interesting than these trends is the departures and approaches to ancestral conditions along different dimensions of the climatic niche exhibited by both ancestral crown-groups and extant major clades (Figs. 6b-c). Some groups maintained similar climatic preferences along the temperature gradient than the ancestral crown group, such as *Elodes-Adenotrias*, *Myriandra*, or *Androsaemum*. Others became adapted to new temperature and precipitation regimes in the subtropical mountain regions, such as Afromontane *Campylosporus*, or Asian *Ascyreia*, but seem to have done so in different directions. The first evolved higher drought tolerance but increased its minimum monthly temperature towards more tropical conditions (similar to Vismieae); the second increased its tolerance to cold temperatures but showed lower tolerance to drought. A similar pattern can be found between *Myriandra* and *Brathys*: both maintained ancestral temperature tolerances but *Brathys* evolved towards wetter climatic conditions, more similar to ancestral nodes 14-15. This heterogeneity or lack of "directionality" in the exploration of the niche space, with bounces back and forth from the intermediate ancestral conditions, can be found also between the ancestral nodes (24 and 25, 23 and 21, Fig. 6b-c), and lends support to Meseguer et al.'s suggestion of clade niche specialization within the genus broad climatic preferences.

Divergence within the ancestral niche did not apparently trigger an increase in diversification rates. MEDUSA only found support for a rate shift towards decreasing rates (lower speciation and lower extinction rates) associated to two clades: the *Androsaemum*-group and especially the early-divergent clade *Elodes-Adenotrias* (Fig. 4). The latter has occasionally been excluded from *Hypericum* on the basis of its specialized floral morphology, with pseudo tubular corollas and vestigial fascicles between the stamens, a character that was found to be plesiomorphic in *Hypericum* (Meseguer et al., 2013). Interestingly, this section retained the ancestral ecological preferences of the ancestor of all living *Hypericum* (crown-group node 18). The *Elodes-Adenotrias* lineage is species-poor (4 species), but also quite old, more than 20 Ma (Meseguer et al., 2013), and is distributed in the North of Africa and Western Palearctic regions. Altogether, these lines of evidence suggest that this lineage, the sister-group of the remaining species (Meseguer et al., 2013), could be a relict of a former western Tethyan group that formed part of the warm-temperate vegetation that dominated the Western Palearctic during the Late Oligocene- Early Miocene (Wolfe, 1975; Raven & Axelrod, 1978). The opposite pattern is found in clade E (*Hypericum-Hirtella*), which shares part of its geographic distribution with *Elodes-Adenotrias*, but it is considerably species-richer and presents very divergent ecological preferences relative to the ancestral conditions. This suggests a different origin, probably a more recent radiation preceding or associated to the onset of the Mediterranean climate (Thompson, 2005).

The only increase in diversification rates detected by TreePar took place in the Late Miocene (9.25) after the LCM event that marked a dramatic decrease in global temperatures. This result does not necessarily contradict the results from MEDUSA. Subclades within the unresolved tips (major clades) could still have experienced changes in diversification rates as suggested by the method-of-moments estimator (Magallón & Sanderson, 2001; Table 1). Improved sampling within these subclades would be necessary to detect these nested rate shifts. Moreover, the increase in diversification rates detected by TreePar is not necessarily related to the LCM event ("climate-driven diversification"), but could be explained

by an increase in the rate of allopatric speciation as a result of expansion into other geographic regions, in particular the subtropical mountains in Africa, South America, and South East Asia. This is supported by fossil-based ENM and biogeographic inference (Meseguer et al., submitted), who found that entrance into the Southern Hemisphere in the Late Miocene coincided with the presence of climatic corridors that allowed lineages to track their temperate niche and move to the south. Indeed, this event is coincident with the entrance through the new uplifted Neogene tropical mountains of *Hypericum* lineages into Africa (*Campylosporus*), South America (*Brathys*), and South East Asia (*Ascyreia*) (Meseguer et al., 2013, submitted). A signal of high allopatric speciation was also found by GeoSSE in these regions, supporting geographic dispersal and expansion.

GeoSSE found no significant association between geography and diversification (i.e., living in a certain region increases or decreases the probability of speciate or go extinct). The only exception is the Eastern Palearctic region (Fig. 5), which shows speciation and extinction rates almost an order of magnitude higher than the other regions. The high extinction and species turnover (ε = 0.88, Table S1) in Asia fits well with Meseguer et al. (submitted) ENM projections, showing a decrease in favourable area for *Hypericum* in this region over time. This might seems surprising since Asia has been postulated as a refuge for temperate lineages during the extreme climatic shifts of the Late Neogene (Sanmartín et al., 2001; Donoghue & Smith 2004). However, the high extinction rates inferred here could refer to the northernmost parts of Asia, while the southern and eastern regions would still have acted as climatic refuges. Two lines of evidences support this: Meseguer et al. ENM projections found that the only part of EP that remained climatically favourable through time was southeastern Asia, and fossil remains of *Hypericum* that are considered a relict assemblage of the mixed-mesophytic forest, have been reported within this region, from Yunnan, China (Meseguer & Sanmartín 2012).

4.3 Drivers of diversification: Key innovation versus Ancestral Resilience?

Along this work we aimed to disentangle the mechanisms that have generated the large species diversity in the angiosperm genus *Hypericum*, in particular the relative contribution of key innovation versus stochastic processes in its origin and diversification. Nonetheless, we found that the answer to this question is not straightforward since different process prevailed at different temporal and spatial scales.

Ecological niche differentiation has been associated with diversification (Moritz et al., 2000; Kozak & Wiens, 2007), but few studies have been able to demonstrate the correlation between the rate at which the niche diverges and the rate at which species accumulates through time (Schnitzler et al., 2012; Kozak & Wiens, 2010).

Along its evolutionary history *Hypericum* experienced periodic episodes of niche change, especially the transition from tropical to temperate habitats experienced at the start of its diversification, but this was not concurrent with change in diversification. The lack of signal might be the result of methodological artefacts as discussed along the text, especially extinction, but it could be also the scale of the study. Some authors have suggested that speciation is rarely the result of niche change (Peterson et al., 1999; Warren et al., 2008) or at least, that this association could only be manifested at fine-scale dimensions (Peterson, 2011). Extinction could have erased the signal in old groups, Early Cenozoic, older than 20 Ma, especially if extinction is not random but affected different clades or geographic regions. Another explanation, however, pointed here is that the large species richness exhibited by *Hypericum* can be explained by the steady accumulation of species over time and that this steadiness is probably rooted in the high genetic plasticity of the ancestral boreotropical lineages, which presented wide tolerances allowing them to cope with climate change. Willis & Macdonald (2011) argued that tropical plants that evolved during the Late Cretaceous under much higher temperature and CO_2 levels presented wider tolerances than it is apparent from their present distributions. This genetic plasticity allows them to persist in the same geographic regions along the major changes of the Early Cenozoic. This seems to be the case for *Hypericum*, whose clusioid ancestors were probably Late Cretaceous tropical plants inhabiting Gondwanan close-canopy forests. The boreotropical ancestors of *Hypericum* seem to have had intermediate climatic preferences with respect to their extant descendants, neither tropical nor temperate (if any *Brathys* seems to be the most similar), which allow them to explore very different dimensions of the niche space, but also gave them a remarkable geographic stability over the dramatic climate changes of the Cenozoic, living to a great extent in the same geographic areas than today extant descendants. Meseguer et al., (submitted) argued that rather than moving into a new ecospace, clade divergence within *Hypericum* was driven by specialization within these intermediate conditions. Indeed, ecological innovation has not been directional in *Hypericum*, but instead unconstrained in the ecological space, probably as the response of very different selective pressures (Fig. 6b-c).

In all, these evidences point out towards the ecological and physiological concept of "resilience". This has been described as the individual's, species', or ecosystem's capacity to cope with disturbance and stress without showing negative effects, which might imply on a bouncing back to a previous state, undergo change while still retaining the same function and capacity, or the

coexistence of multiple stable states or regimes (Holling, 1973; Sgro et al., 2010). This definition fits well the evolutionary patterns found in *Hypericum*, which show a bouncing back to ancestral climatic preferences as well as exploration of different dimensions of the niche space, a remarkable degree of geographic stability (i.e., the ability to persist in the same regions without getting extinct), but also the capacity to evolve new adaptations while still retaining the ancestral characteristics (the herbaceous and woody habits). Willis & MacDonald (2011) suggested that the ancestral resilience or plasticity exhibited by Early Eocene plants might have been lost during the course of the Cenozoic through genetic specialization and phylogenetic splits. The implications of this hypothesis are that the ability of an organism to evolve niche preferences to climate change is reduced through time, with range shifts dominating over evolutionary novelties later in the Cenozoic. Our results, however, suggest that present clades have the same capacity to explore very different dimensions of the niche space than their ancestors (Fig. 6b-c). This has interesting implications, as it might indicate that current *Hypericum* species have the same ability to cope with climate change as their ancestors.

Much interest has been focused on the study of adaptive radiations and key innovations to understand the generation of biodiversity, usually dealing with young, recently diverged groups. However, results from this work demonstrate that ancient lineages like *Hypericum*, which have survived through the climate changes of the Cenozoic, also merit attention for what they can tell us about the role of "resilience" and "plasticity" in coping with extinction. Like a *marathon runner*, as opposed to a "sprinter", *Hypericum* has been able to survive and diversify keeping a steady pace through all major climatic changes of the Cenozoic, allowing it to reach a large diversity and wide geographic distribution.

References

Ackerly DD (2004) Adaptation, niche conservatism, and convergence: Comparative studies of leaf evolution in the California chaparral. Am Nat 163:654–671.

Alfaro ME, Santini F, Brock C, Alamillo H, Dornburg A, Rabosky DL, Carnevale G, Harmon LJ (2009) Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. Proc Natl Acad Sci USA 106:13410–13414.

Antonelli AA, Sanmartín I (2011). Mass extinction, gradual cooling, or rapidradiation? Reconstructing the spatiotemporal evolution of the ancient angiosperm genus *Hedyosmum* (Chloranthaceae) using empirical and simulated approaches. Systematic Biology 60: 596–615.

Arbuzova O, Chelebaeva A, Iljinskaja I, Proskurin K, Vickulin S (2005) In: Budantsev, L. & Iljinskaja, I. (eds.), Fossil flowering plants of Russia and adjacent countries, Vol. 4 Nyctaginaceae-Salicaceae. Izdatelstvo Nauka Leningradskoe otd-nie, 1974-. Leningrad.

Baldwin BG, Sanderson MJ (1998) Age and rate of diversification of the Hawaiian silversword alliance (Compositae). Proc Natl Acad Sci USA 95:9402–9406.

Benton MJ (2009) The Red Queen and the Court Jester: species diversity and the role of biotic and abiotic factors through time. Science 323: 728-732.

Beerling DJ et al. (2012) Ecosystem CO_2 starvation and terrestrial silicate weathering: mechanisms and global-scale quantification during the late Miocene. J Ecol 100: 3141.

Beerling DJ, Foxa A, Stevenson DS, Valdes PJ (2011) Enhanced chemistry-climate feedbacks in past greenhouse worlds. Proc Natl Acad Sci USA 108: 9770–9775.

Beerling DJ, Royer DL (2011) Convergent Cenozoic CO2 history. Nature 4: 418–420

Beerling D, Berner RA, Mackenzie FT, Harfoot M, Pyle JA (2009) Methane and the CH4-related greenhouse effect over the past 400 million years. Am J Sci 309:97–113.

Bradshaw CD et al. (2012) The relative roles of CO_2 and palaeogeography in determining late Miocene climate: results from a terrestrial model-data comparison. Clim Past 8: 1257–1285.

Brown JW, FitzJohn RG, Alfaro ME, Harmon LJ (2012) MEDUSA: Modeling Evolutionary Diversification Using Stepwise AIC. Available from: https://github.com/josephwb/turboMEDUSA.

Calenge C, Basille M (2008) A general framework for the statistical exploration of the ecological niche. J Theor Biol 252:674–685.

Carine MA, Christenhusz MJM (2010) About this volume: the monograph of *Hypericum* by Norman Robson. Phytotaxa 4: 1–4.

Cavagnetto C, Anadón P (1996) Preliminary palynological data on floristic and climatic changes during the Middle Eocene-Early Oligocene of the eastern Ebro Basin, northeast Spain. Review of Palaeobotany and Palynology 92: 281-305.

Cracraft J (1990) The origin of evolutionary novel pattern and process at different hierarchical levels. Evolutionary innovations. pp. 21–46

Crepet WL, Nixon KC (1998) Fossil Clusiaceae from the late Cretaceous (Turonian) of New Jersey and implications regarding the history of bee pollination. American Journal of Botany 85: 1122-1133.

Crisp M, Arroyo MTK, Cook LG, Gandolfo MA, Jordan GJ, McGlone MS, Weston PH, Westoby M, Wilf P, Linder HP (2009) Phylogenetic biome conservatism on a global scale. Nature 458: 754-756.

Davis CC, Webb CO, Wurdack KJ, Jaramillo CA, Donoghue MJ (2005) Explosive radiation of Malpighiales supports a mid-Cretaceous origin of modern tropical rain forests. Amer Nat 165: E36-E65.

Donoghue MJ (2008) A phylogenetic perspective on the distribution of plant diversity. Proc Natl Acad Sci USA 105:11549-11555.

Donoghue MJ, Moore BR (2003) Toward an integrative historical biogeography. Journal of Integrative and Comparative Biology 43:261–270.

Eldredge N, Thompson JN, Brakefield PM, Gavrilets S, Jablonski D, Jackson JBC et al. (2005). The dynamics of evolutionary stasis. Paleobiology 31:133–145.

Evans ME, Smith SA, Flynn RS, Donoghue MJ (2009) Climate, niche evolution, and diversification of the "bird-cage" evening primroses (Oenothera, sections Anogra and Kleinia). The American Naturalist 173: 225–240.

Ezard THG, Aze T, Pearson PN, Purvis A (2011) Interplay between changing climate and species ecology drives macroevolutionary dynamics. Science 332: 349-351.

FitzJohn RG, Maddison WP, Otto SP (2009) Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. Syst Biol 58: 595–611.

FitzJohn RG (2010) Quantitative traits and diversification. Syst. Biol. 59:619–633.

Graham CH, Ron SR, Santos JC, Schneider CJ, Moritz C (2004) Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. Evolution. 58:1781–1793.

Goldberg EE, Lancaster LT, Ree RH (2011). Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. Syst Biol 60: 451–465.

Heard SB, Hauser DL (1995) Key evolutionary innovations and their ecological mechanisms. Historical Biology 10 (2): 151–173.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. Int J Climatol 25:1965–1978.

Hirzel AH, Hausser J, Chessel D, Perrin N (2002) Ecological-niche factors analysis: how to compute habitat-suitability maps without absence data? Ecology 83: 2027–2036.

Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W (2008) GEIGER: investigating evolutionary radiations. Bioinformatics 24:129–131.

Hodges SA (1997) Floral nectar spurs and diversification. International Journal of Plant Sciences 158(suppl.):S81–S88.

Holling CS (1973) Resilience and stability of ecological systems. Annual Review of Ecology and Systematics 4: 1–2

Holt RD (1990). Microevolutionary consequences of climate change. Trends Ecol Evol 9: 311–315.

Holt RD (2009). Bringing the Hutchinsonian niche into the 21st century: ecological and evolutionary perspectives. Proc Natl Acad Sci USA 106: 19659–19665.

Holt RD, Gaines MS (1992). Analysis of adaptation in heterogeneous landscapes: implications for the evolution of fundamental niches. Evol Ecol 6: 433–447.

Hughes C, Eastwood R (2006) Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. Proceedings of the National Academy of Sciences of the USA 103:10334–10339.

Jablonski D, Roy K, Valentine JW (2006) Out of the tropics: evolutionary dynamics of the latitudinal diversity gradient. Science 314:102–106.

Kottek M, Grieser J, Beck C, Rudolf B, Rubel F (2006) World Map of the Köppen-Geiger climate classification updated. Meteorol Z 15: 259-263.

Kozak KH, Wiens JJ (2010) Accelerated rates of climatic-niche evolution underlie rapid species diversification. Ecology Letters 13: 1378–1389.

Leslie AB, Beaulieu JM, Raic HS, Cranea PR, Donoghue MJ, Mathews S (2012). Hemisphere-scale differences in conifer evolutionary dynamics. Proc Natl Acad Sci USA 109:16219.

Lomolino MV, Riddle BR, Brown JH. 2005. Biogeography. Sunderland, MA: Sinauer Associates Inc. 845 pp. 3rd ed.

Lovette IJ, Bermingham E, Ricklefs RE (2002) Cladespecific morphological diversification and adaptive radiation in Hawaiian songbirds. Proceedings of the Royal Society B: Biological Sciences 269: 37–42.

Maddison WP, Midford PE, Otto SP (2007) Estimating a binary character's effect on speciation and extinction. Syst Biol 56: 701–710.

Magallon S, Sanderson MJ (2001) Absolute diversification rates in angiosperm clades. Evolution 55: 1762–1780.

McKnight TL, Hess D (2000) Climate Zones and Types: Highland Climate (Zone H). Physical Geography: A Landscape Appreciation. Upper Saddle River, NJ: Prentice Hall pp 237–40

Meseguer AS, Lobo JM, Ree R, Beerling DJ, Sanmartín I. Using fossils to reveal the impact of ancient climate change in plant evolution. Submitted PLoSBiology.

Meseguer AS, Aldasoro JJ, Sanmartín I (2013) Bayesian inference of phylogeny, morphology and range evolution reveals a complex evolutionary history in St John's wort (Hypericum). Mol Phylo Evol 67(2):379-403.

Meseguer AS, Sanmartín I (2012) Paleobiology of the genus Hypericum (Hypericaceae): A survey of the fossil record and its palaeogeographic implications. An Jar Bot Mad 69:97-106.

Morlon H, Parsons TL, Plotkin J (2011) Reconciling molecular phylogenies with the fossil record. Proc Natl Acad Sci USA 108: 16327–16332.

Moritz C, Patton JL, Schneider CJ, Smith TB (2000). Diversification of rainforest faunas: an integrated molecular approach. Annu Rev Ecol Syst 31: 533–563.

Nee S, Holmes EC, May RM, Harvey PH (1994) Extinctions rates can be estimated from molecular phylogenies. Philos Trans R Soc Lond 344: 77–82.

Nee S, May RM, Harvey PH (1994) The reconstructed evolutionary process. Philos Trans R Soc Lond B Biol Sci 344: 305–311.

Nürk NM, Madriñán S, Carine MA, Chase MW, Blattner FR (2013) Molecular phylogenetics and morphological evolution of St. John's wort (*Hypericum*; Hypericaceae). Mol Phylogenet Evol 66:1-16.

Nürk NM (2011) Phylogenetic analyses in St. John's wort (*Hypericum*). Inferring character evolution and historical biogeography. Dissertation thesis, Free University of Berlin, Berlin, Germany.

Peterson AT (2011) Ecological niche conservatism: a time structured review of evidence. J Biogeogr 28:817-27.

Peterson AT, Soberón J, Sánchez-Cordero V (1999) Conservatism of ecological niches in evolutionary time. Science 285: 1265–1267.

Plana, V., 2004. Mechanisms and tempo of evolution in the African Guineo-Congolian rainforest. Philosophical Transactions of the Royal Society B: Biological Sciences 359, 1585–1594.

Pybus OG, Harvey PH (2000) Testing macro-evolutionary models using incomplete molecular phylogenies. Proc R Soc Lond B 267:2267–2272.

Quental TB, Marshall CR (2010) Diversity dynamics: molecular phylogenies need the fossil record. Trends Ecol Evol 25:434–441.

Rabosky DL (2006a) LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates. Evol Bioinform Online 2:257–260.

Rabosky DL (2006b) Likelihood methods for detecting temporal shifts in diversification rates. Evolution 60:1152–1164.

Rabosky DL (2009) Ecological limits and diversifi cation rate: Alternative paradigms to explain the variation in species richness among clades and regions. Ecology Letters 12: 735 – 743 .

Ree, R. H. 2005. Detecting the historical signature of key innovations using stochastic models of character evolution and cladogenesis. Evolution 59:257–265.

Ricklefs RE (1987) Community diversity: relative roles of local and regional processes. Science 235: 167–171.

Robson NKB (1977) Studies in the genus *Hypericum* L. (Guttiferae). 1. Infrageneric classification. Bulletin of the British Museum of Natural History (Botany) 5: 295–355.

Robson NKB (1981) Studies in the genus *Hypericum* L. (Guttiferae). 2. Characters of the genus. Bulletin of the British Museum of Natural History (Botany) 8: 55–226.

Robson NKB (2012) Studies in the genus Hypericum L. (Hypericaceae) 9. Addenda, corrigenda, keys, lists and general discussion. Phytotaxa 72: 1–111.

Ronquist F, Sanmartín I (2011) Phylogenetic methods in historical biogeography. Ann Rev Ecol Evol Syst 42:441-464.

Rosenzweig ML (1995) Species diversity in space and time. Cambridge, UK: Cambridge University Press.

Ruhfel BR et al (2011) Phylogeny of the clusioid clade (Malpighiales): Evidence from the plastid and mitochondrial genomes. Am J Bot 98:306-325.

Ruhfel BR (2011) Systematics and biogeography of the clusioid clade (Malpighiales). Dissertation thesis. Harvard University, Cambridge, Massachusetts.

Sanmartín I, Engho_ H, Ronquist F (2001) Patterns of animal dispersal, vicariance and diversification in the Holarctic. Biol J Linnean Soc 73:345-390.

Sargent RD (2004) Floral symmetry affects speciation rates in angiosperms. Proceedings of the Royal Society B: Biological Sciences 271:603–608.

Savolainen V, Heard SB, Powell MP, Davies TJ, MOOE AØ (2002) Is Cladogenesis Heritable? Syst Biol 51(6):835–843.

Schluter D, Price T, Mooers A, Ludwig D (1997) Likelihood of ancestor states in adaptive radiation. Evolution. 51:1699–1711.

Schnitzler J, Graham CH, Dormann CF,Schi_ers K, Linder HP (2012) Climatic niche evolution and species diversi_cation in the Cape flora, South Africa. J Biogeogr 39:22012211.

Sgro CM, Lowe AJ, Hoffman AA (2010). Building evolutionary resilience for conserving biodiversity under climate change. Evol Appl 4: 326-337.

Smith SA, Donoghue MJ (2010) Combining Historical Biogeography with Niche Modeling in the Caprifolium Clade of Lonicera (Caprifoliaceae, Dipsacales). Syst Biol 59:1-20.

Smith SA, Beaulieu JM (2009). Life history influences rates of climatic niche evolution in flowering plants. Proc R Soc Lond B 276: 4345–4352.

Schneck R, Micheels A, Mosbrugger V (2012) Climate impact of high northern vegetation: Late Miocene and present. Int J Earth Sci (Geol Rundsch) 101:323338. DOI10.1007/s00531-011-0652-4.

Schnitzler J, Graham CH, Dormann CF,Schi_ers K, Linder HP (2012) Climatic niche evolution and species diversi_cation in the Cape flora, South Africa. J Biogeogr 39:2201-2211.

Stadler T (2011). Mammalian phylogeny reveals recent diversification rate shifts. Proc Natl Acad Sci USA 108: 6187–6192.

Stevens PF (2007) Hypericaceae. In: The families and genera of vascular plants Vol. 9, Kubitzki K (eds). Springer, Berlin, pp. 194–201.

Tiffney BH (1985a) Perspectives on the origin of the floristic similarity between eastern Asia and eastern North America. J Arnold Arbor 66:73–94.

Tiffney BH (1985b) The Eocene North Atlantic land bridge: its importance in Tertiary and modern phytogeography of the Northern Hemisphere. J Arnold Arbor 66:243-273.

Thompson JD (2005) Plant evolution in the Mediterranean. Oxford University Press.

Valencia-Barrera RM, Comtois P, Fernández-González D (2002) Bioclimatic indices as a tool in pollen forecasting. Int J Biometeoro l46:171–175.

Vamosi JC, Vamosi SM (2010) Factors influencing diversification in angiosperms: at the crossroads of intrinsic and extrinsic traits 1. American Journal of Botany 98(3): 460–471. 2011.

Varela S, Lobo JM, Hortal J (2011) Using species distribution models in paleobiogeography: A matter of data, predictors and concepts. . Palaeogeogr Palaeoclimatol Palaeoecol 310:451-463.

Vieites, D.R., Nieto-Roma´n, S. & Wake, D.B. (2009) Reconstruction of the climate envelopes of salamanders and their evolution through time. Proceedings of the National Academy of Sciences USA, 106, 19715–19722.

Warren DL, Glor RE, Turelli M (2008). Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution 62: 2868–2883.

Waser N (1998) Pollination, angiosperm speciation, and the nature of species boundaries. Oikos 81:198–201.

Weir JT, Schluter D (2007) The latitudinal gradient in recent speciation and extinction rates of birds and mammals. Science 315: 1574–1576.

Wenn (1999) Evolution of the eastern Asia and eastern North American disjunct distributions in flowering plants. Annu Rev Ecol Syst 30:421–55.

Wiens JJ, Donoghue MJ (2004) Historical biogeography, ecology and species richness. Trends Ecol Evol 19:639-644.

Wiens JJ, Ackerly DD, Allen AP, Anacker BL, Buckley LB, Cornell HV, Damschen EI, Davies TJ, Grytnes JA, Harrison SP, Hawkins BA, Holt RD, McCain CM, Stephens PR (2010) Niche conservatism as an emerging principle in ecology and conservation biology. Ecology Letters 13: 1310–1324.

Wiens JJ (2007) Global patterns of diversification and species richness in amphibians. Am Nat 170:S86–S106.

Willis KJ, MacDonald GM (2011). Long-term ecological records and their relevance to climate change predictions for a warmer world. Annu Rev Ecol Evol Syst 42:267-87.

Wolfe JA (1969) Neogene floristic and vegetational history of the Pacific North-west. Madroño 20: 83-110.

Wolfe JA (1972) An interpretation of Alaskan Tertiary floras. In: Graham A (eds) Floristics and Paleofloristics of Asia and eastern North America. Elsevier, Amsterdam: 201-233.

Wolfe JA (1975) Some aspects of plant geography of the northern hemisphere during the Late cretaceous and Tertiary. Annals of the Missouri Botanical Garden 62: 264-279.

Xiang QY, Soltis DE (2001) Dispersal-vicariance analyses of intercontinental disjuncts: Historical biogeographical implications for angiosperms in the Northern Hemisphere. Int J Plant Sci 162:S29–39.

Xiang QY, Soltis DE, Soltis PS (1998) The eastern Asian and eastern and western North America floristic disjunction: Congruent phylogenetic patterns in seven diverse genera. Mol Phylo Evol 10:178–190.

Xiang QY, Soltis DE, Soltis PS, Manchester SR, Crawford DJ (2000) Timing the eastern Asian–eastern North American floristic disjunction: Molecular clock corroborates paleontological estimates. Mol Phylo Evol 15:462–472.

Yesson C, Culham A (2006) Phyloclimatic modeling: combining phylogenetics and bioclimatic modeling. Syst Biol 55:785-802.

Zachos JC, Dickens GR, Zeebe RE (2008). An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. Nature 451: 279–283.

TABLES

Table 1. Species richness and absolute rate of diversification for major lineages within *Hypericum* and the clusioid clade. Rates of diversification were estimated in absence of extinction (ε = 0) and under a high relative extinction rate (ε = 0.99) for both the stem group and crown group age if possible. Rates of diversification are presented in ascending order according to the value of ε = 0 in the stem age, which is the most abundant value. Rates of inclusive clades (Hypericaceae, genus *Hypericum*) are in bold. Moment estimator: probability values indicating departure of lineages from the global rate of diversification for *Hypericum* (rH) and the clusioid clade (rC) estimated using method-of-moments estimator (Magallón & Sanderson, 2001); significant values are marked with an asterisk.

Clade	# spp.	Crown	Stem	Rate estimate stem age		Rate estimate crown age		Moment estimator	
		age	**age**	**ε = 0**	**ε = 0.99**	**ε = 0**	**ε = 0.99**	**ε = 0**	**ε = 0.99**
genus *Hypericum*	**496**	**34.92**	**49.9**	**0.1244**	**0.0357**	**0.15788**	**0.05092**	**rH=0.1244**	**rH=0.0357**
Webbia	1	-	10.84	0	0	-	-	1	1
Campylopus	1	-	15.78	0	0	-	-	1	1
Elodes	1	-	27.72	0	0	-	-	1	1
Adenotrias	3	-	27.72	0.0146	0.00048	-	-	0.9374	0.3947
Triadenioides	3	-	21.02	0.01928	0.0006399	-	-	0.859	0.2786
Triadenum	8	-	29.3	0.0473	0.0021	-	-	0.8308	0.0483
Arthrophyllum	5	-	17.29	0.05299	0.00194	-	-	0.6096	0.0449
Androsaemum	4	5.21	10.84	0.0639	0.00218	0.133	0.0045	0.4058	0.0330*
Campylosporus	7	4.01	15.99	0.0783	0.0033	0.3124	0.01317	0.4136	0.0067*
Myriandra	29	13.59	21.92	0.12	0.011	0.196	0.0177	0.1504	3.71E-08*
Hirtella	71	14.84	18.85	0.1893	0.02788	0.24053	0.03541	0.0009*	2.00E-22*
Brathys	139	13.68	21.92	0.193	0.039	0.31	0.063	8.81E-05*	2.38E-37*
Ascyreia	56	12.33	15.99	0.2083	0.02709	0.27025	0.0351	0.0003*	1.30E-20*
Hypericum	139	13.31	15.78	0.268	0.0546	0.31865	0.06476	8.52E-10*	3.28E-51*
Clusioids	**1900**	**91.68**	**-**	**-**	**-**	**0.082347**	**0.03267051**	**rC=0.082347**	**rC= 0.0327**
Terniopsis	6	21.14	52.68	0.034012	0.000926	0.0520	0.0165	0.3813	0.0309*
Hypericum	496	34.84	57.11	0.1076975	0.0304141	0.1566	0.1095	2.75E-13*	9.35E-84*
Bonnetia	30	16.1	40.55	0.083876	0.006279	0.1682	0.0813	0.0001*	5.51E-12*
Psorospermum	48	19.66	35.25	0.10982	0.010929	0.1969	0.019596	3.11E-05*	5.73E-16*
Vismia	55	17.65	28.37	0.141252	0.015219	0.22704	0.0244635	5.69E-07*	4.48E-20*
Garcinia	260	18.3	45.76	0.1215184	0.027931	0.2660	0.1771	6.66E-29*	1.54E-90*
Harungana	50	11.6	28.37	0.1378	0.01405	0.33724	0.034377	4.62E-11*	2.80E-25*
Clusia	300	6.24	10.03	0.568672	0.137965	0.8030	0.5416	4.00E-119*	3.03E-220*
Calophyllum	186	4.89	9.94	0.52572	0.105364	0.9269	0.5969	1.92E-89*	2.03E-154*
Weddellina	1	60.17	60.17	0	0	NA	NA	1	1
Endodesmia	1	45.04	45.04	0	0	NA	NA	1	1

Table 2. Results of the TreePar analysis for the BEAST MCC of *Hypericum* shown in Fig. 2. Maximum likelihood parameter estimates for 2 shifts. Abbreviations: LH: log likelihood value; ε_i is the turnover (extinction/speciation) estimates for the successive intervals going back in time; r_i, is the diversification rate estimate (speciation-extinction) for the same intervals. Time shift i, are the shift times for those intervals; LRT: probability value for the Likelihood ratio test (LRT).

Model	**Rate changes**	**LH**	**ε_1 present**	**$\varepsilon_{2\ past}$**	**$\varepsilon_{3\ past}$**	**r_1**	**r_2**	**r_3**	**Time shift 1**	**Time shift 2**	**LRT**
1	0 shift	260.70	0.69	-	-	0.146	-	-	-	-	-
100%	**1 shift**	**256.5**	**0.49**	**0.93**	**-**	**0.197**	**0.0457**	**-**	**9.25**	**-**	**0.038***
surviv	2 shift	254.4	0.557	0.999	0.989	0.183	0.0002	0.053	9.25	27.07	0.257
2	0 shift	260.70	0.69	-	-	0.146	-	-	-	-	-
10%	**1 shift**	**256.5**	**0.49**	**0.993**	**-**	**0.197**	**0.046**	**-**	**9.25**	**-**	**0.038***
surviv	2 shift	255.13	0.181	1.091	0.999	0.261	-0.28	0.046	7.27	9.25	0.432

Table 3. Ancestral niche reconstruction for maximum and minimum values of the selected continuous variables performed in the R package *Geiger* using maximum likelihood under a Brownian model. Node numbers correspond with Figure 6. Abbreviations: "Annual Prec" Annual precipitation, "Annual Var Prec" Annual variation in precipitation, "Max Month Prec" Maximum month precipitation, "Min Month Prec" Minimum month precipitation, "Min Month Temp" Minimum month temperature.

Node number	Annual Prec max	Annual Prec min	Annual Var Prec max	Annual Var Prec min	Aridity max	Aridity min	Continentality max	Continentality min	Max Month Prec max	Max Month Prec min	Min Month Prec max	Min Month Prec min	Min Month Temp max	Min Month Temp min
14	2890.376	354.9077	355.0569	54.29509	50.6273	6.42247	24.73806	4.313777	386.1388	66.57202	139.3708	1.396956	22.61024	-2.34379
15	2790.08	330.9824	353.2156	51.95456	49.04952	6.010724	25.3093	4.558408	381.2029	62.92313	129.8750	1.426260	22.28618	-3.10739
16	2970.126	335.5015	405.5195	73.13823	50.70769	6.00905	15.5433	2.557877	415.9177	77.99596	113.1070	0.422620	24.59629	5.109154
17	3186.825	269.5815	423.8548	72.02493	53.96443	4.983534	12.45578	1.838968	430.2447	75.2804	120.1698	0.135417	25.55022	7.753283
18	2305.732	238.2407	315.2574	30.57798	42.1267	4.461794	33.25282	6.665982	342.0301	40.2465	104.0775	2.131930	19.69597	-10.85588
19	2311.72	230.7251	318.7842	29.08605	42.26063	4.321947	34.1887	6.685614	345.4875	38.67494	103.7708	1.956041	19.81902	-11.76307
20	2308.505	256.2063	322.7973	28.97579	42.18574	4.86256	36.18266	6.906556	351.9818	40.92436	99.21646	1.903344	19.84721	-13.96396
21	2435.156	247.8809	344.9928	26.77413	44.12142	4.716785	36.51129	4.732731	372.2938	39.34648	101.5418	1.540135	22.12004	-15.58673
22	2350.014	150.9288	332.5949	20.89462	43.12149	2.73875	36.47865	6.468566	355.2375	26.54534	108.7932	1.049065	20.46638	-13.58607
23	2416.691	130.6612	337.6382	18.97257	44.20791	2.353013	37.99771	5.916094	358.2345	23.54549	114.4438	0.775152	20.88742	-15.22549
24	2934.657	106.4945	340.1295	18.73041	51.27535	1.659693	36.71781	3.909382	362.9824	21.47247	149.8301	0.377943	23.50724	-13.31645
25	2208.742	105.9819	346.1021	15.28053	41.72451	2.037211	41.88587	6.132268	361.1095	18.9009	102.4325	0.488175	20.00355	-19.76976

Table 4. Ancestral niche reconstruction for the tolerance of selected continuous variables performed in *geiger*. Node numbers correspond with Figure 6.

Node number	Tolerance Annual Prec	Tolerance Annual Var Prec	Tolerance Aridity	Tolerance Continentality	Tolerance Max Month Prec	Tolerance Min Month Prec	Tolerance Min Month Temp
14	2535.469	300.7618	44.20483	20.42428	319.5667	137.97384	24.95404
15	2459.097	301.261	43.0388	20.75089	318.2798	128.44883	25.39358
16	2634.624	332.3812	44.69864	12.98542	337.9217	112.68447	19.48713
17	2917.244	351.8299	48.9809	10.61681	354.9643	120.03445	17.79693
18	2067.491	284.6794	37.66491	26.58684	301.7836	101.9456	30.55185
19	2080.995	289.6981	37.93869	27.50309	306.8126	101.81476	31.58209
20	2052.298	293.8215	37.32318	29.2761	311.0574	97.31312	33.81117
21	2187.275	318.2186	39.40464	31.77856	332.9473	100.00172	37.70678
22	2199.085	311.7002	40.38274	30.01009	328.6922	107.74414	34.05246
23	2286.03	318.6656	41.8549	32.08161	334.689	113.66873	36.11292
24	2828.162	321.3991	49.61566	32.80843	341.5099	149.45216	36.8237
25	2102.76	330.8216	39.6873	35.7536	342.2086	101.94433	39.77331

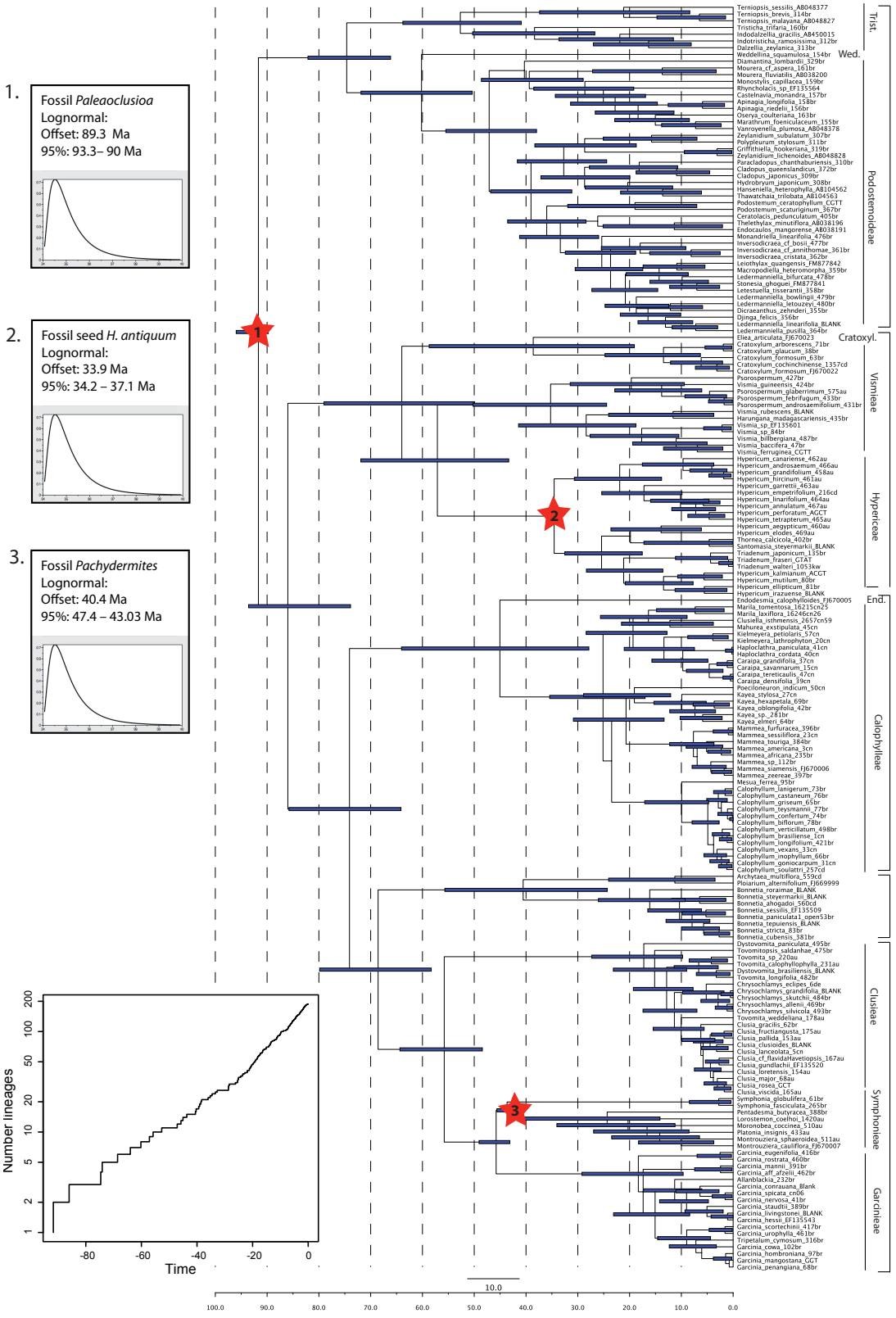


FIGURE 1. Divergence time estimation analysis of the clusioid clade in BEAST, showing median divergence times and 95% confidence intervals (for main lineages), derived from Ruhfel et al. (2011) multigene phylogeny. Divergence time estimates were obtained by using three fossil constraints showed by a red start. Scale bar represents the major Cretaceous and Cenozoic intervals. Inset figure: Lineage through time (LTT) plot with the MCC tree. Trist. = Tristichoideae , Wed.= Weddellinoideae, Cratoxyl. = Cratoxyleae, End. = Endodesmieae.

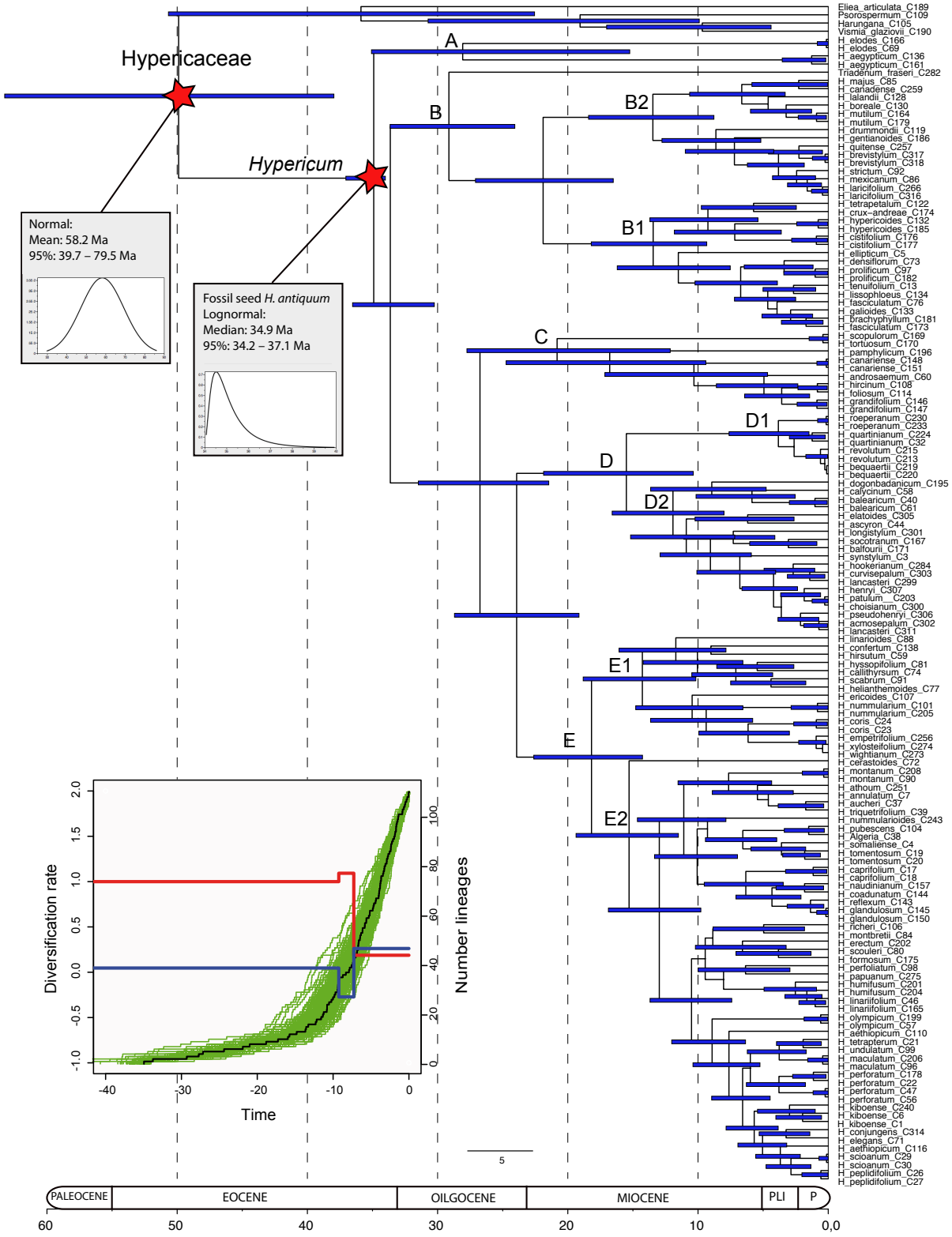


Figure 2. Divergence time estimation analysis of *Hypericum* in BEAST showing median divergence times and 95% confidence intervals (for main lineages), derived from the ‘‘Two-marker’’ concatenate dataset of Meseguer et al. (2013). Inset figure: Lineage through time (LTT) plot for 100 trees of the posterior sample of BEAST in green, with the MCC tree in black. Net diversification (speciation – extinction) rate (blue line) and species turn over (speciation/extinction; in red) calculated in 0.34 Ma intervals and allowing a maximum of 8 rate shifts. Sampling fraction was fixed to 0.1 (only 10% of the species survive from one period to the other, see text). Clade numbers correspond to Meseguer et al. (2013).

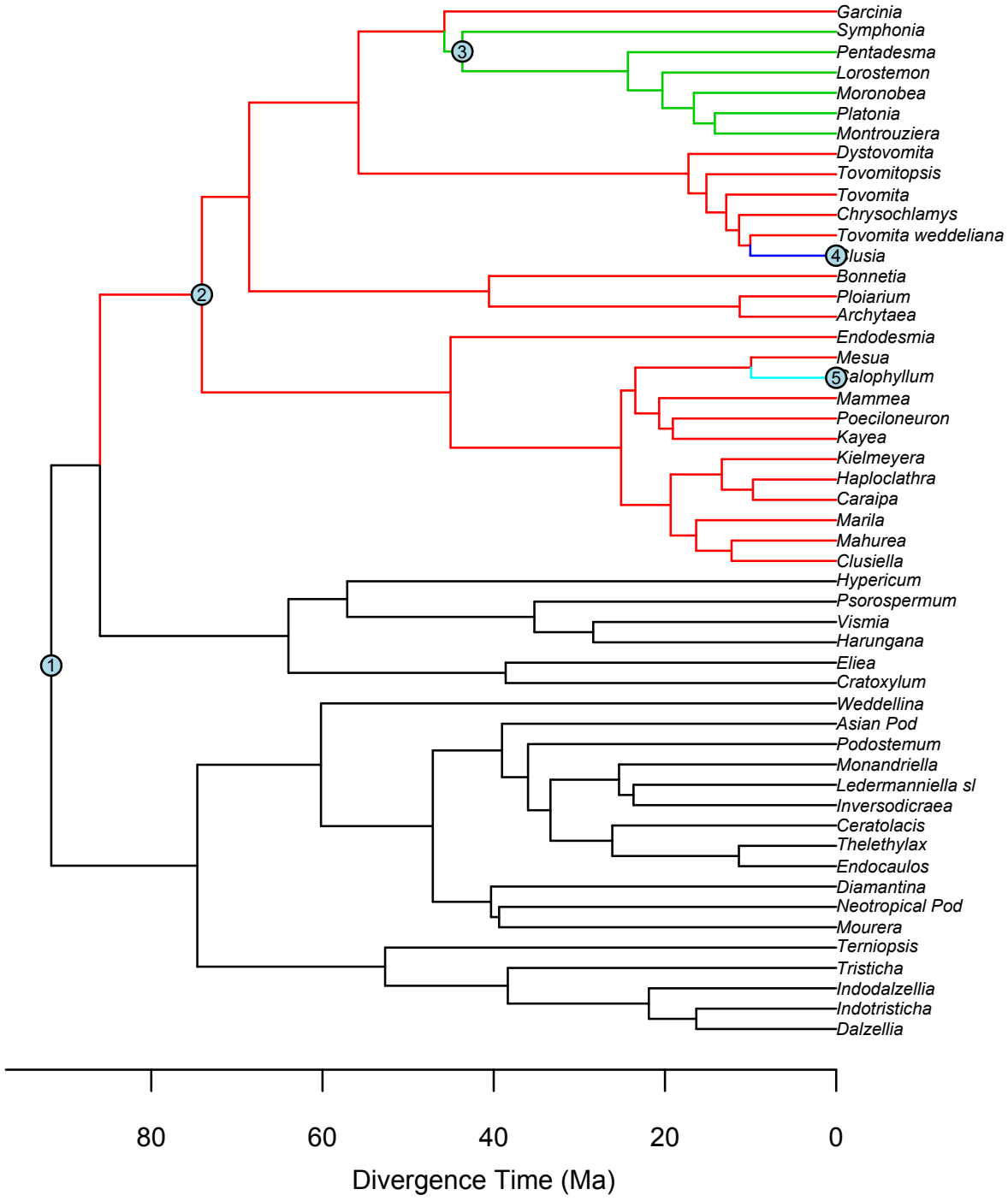


Figure 3. Pruned chronogram of the clusioid clade showing major lineages. Colour branches indicate clades that present different diversification rates that the background rate of the tree, as identified by the TurboMedusa model.

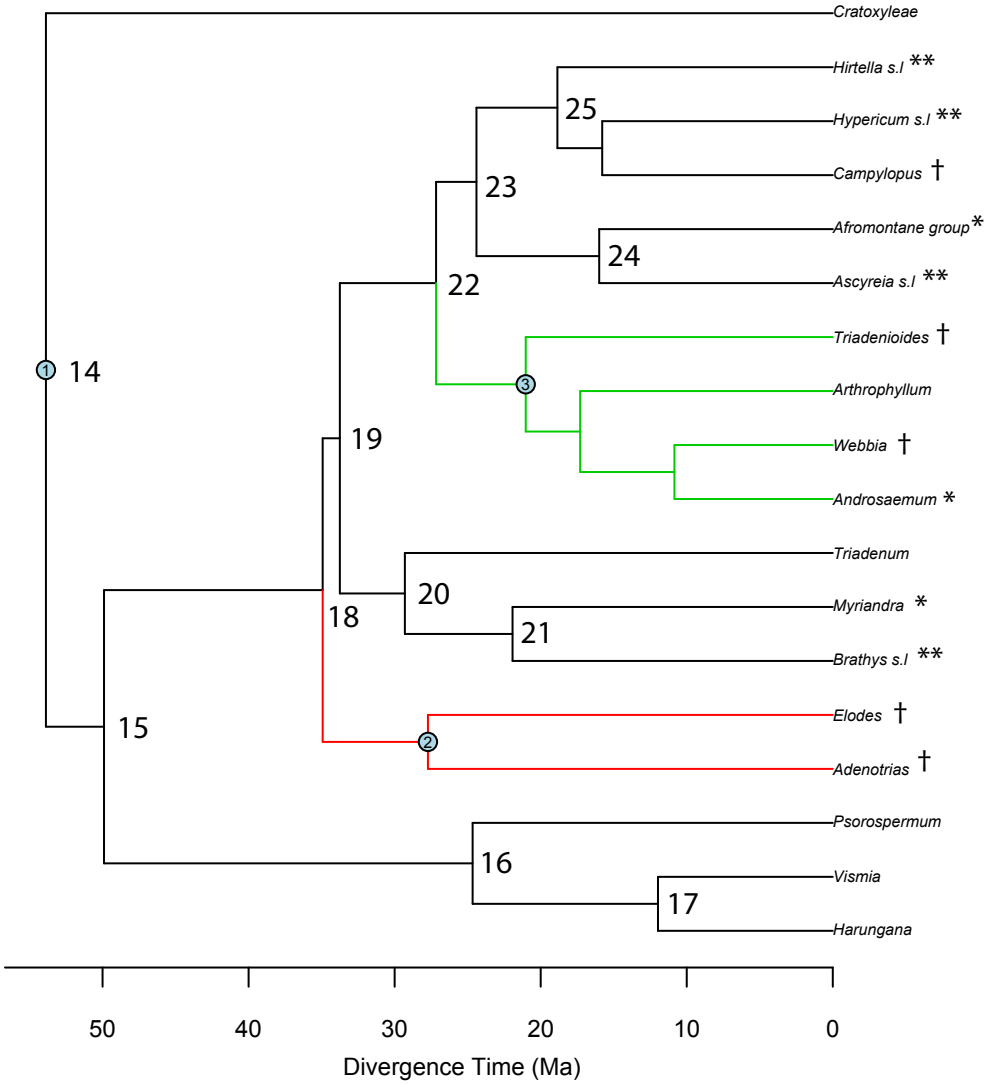


Figure 4. Pruned phylogeny of *Hypericum* showing major lineages as recovered by Meseguer et al. (2013). Colour branches indicate clades that present different diversification rates that the background rate of the tree, as identified by the TurboMedusa model. The symbol "*" highlights those clades that present significantly higher species richness than expected, given its age and a very high extinction rate, under the global rate of cladogenesis of *Hypericum*, as resulted from the moment estimator (Magallón & Sanderson, 2001). The symbol "**" indicates higher species richness than expected under both very low and very high extinction values (see table X). The symbol "†" is subtending to clades that present significantly lower species richness than expected under the global rate of cladogenesis. Node numbers are indicated.

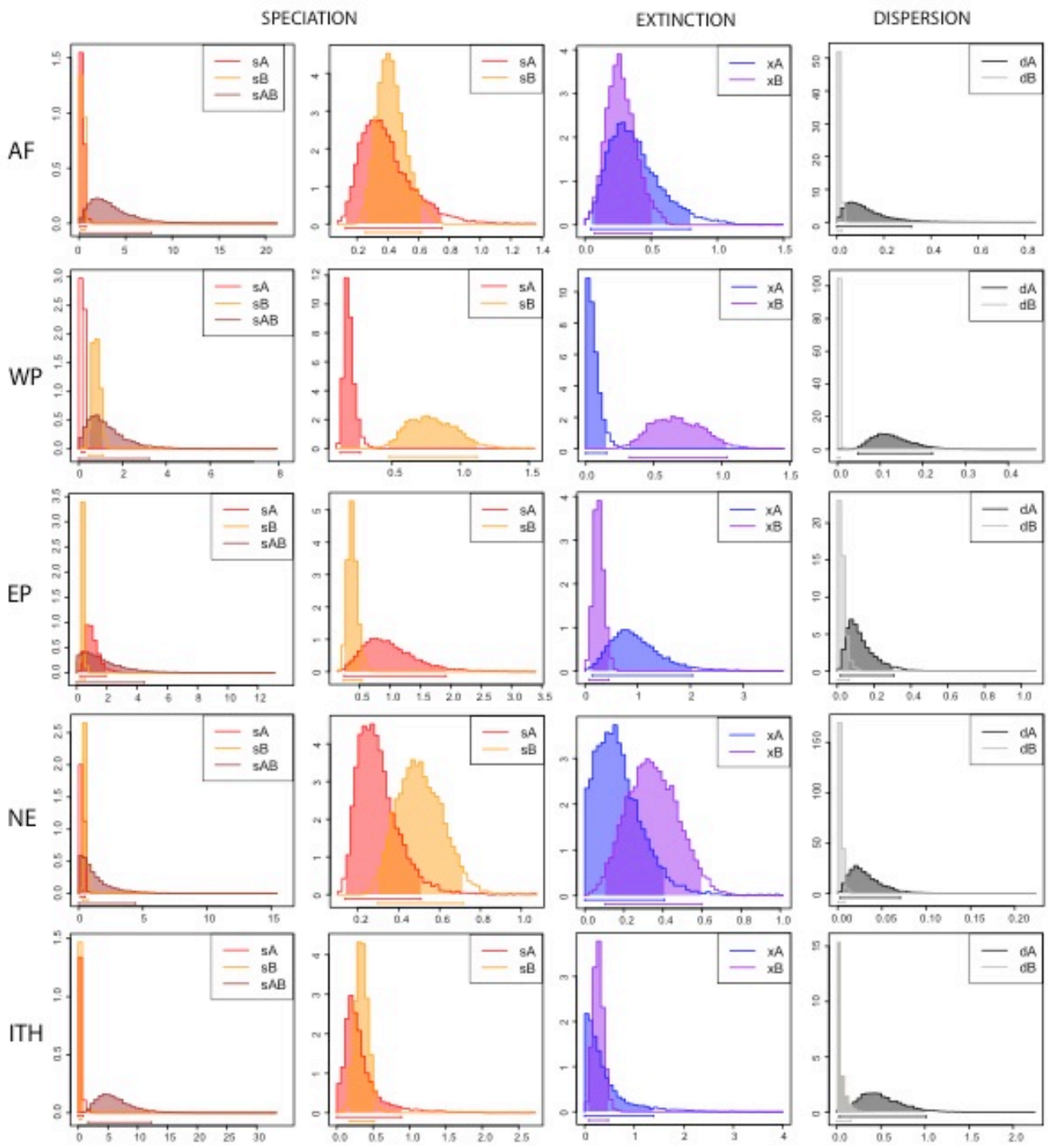


Figure 5. Posterior probability distributions for the speciation and extinction parameters across biogeographic regions. Speciation and extinction shown within a given region (sA and xA, respectively) and the remainder of the distributional range of Hypericum (sB and xB, respectively) for each area estimated by the GeoSSE (see text for more details). Biogeographic regions correspond with the areas in the map.

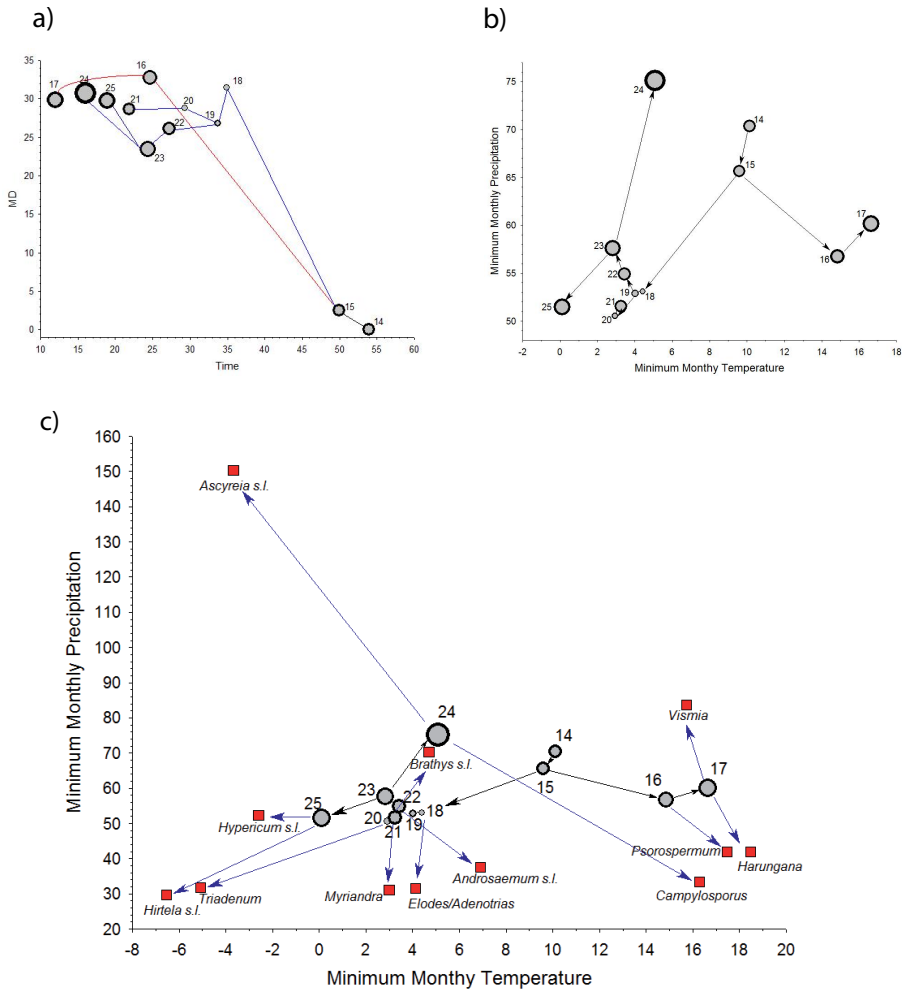
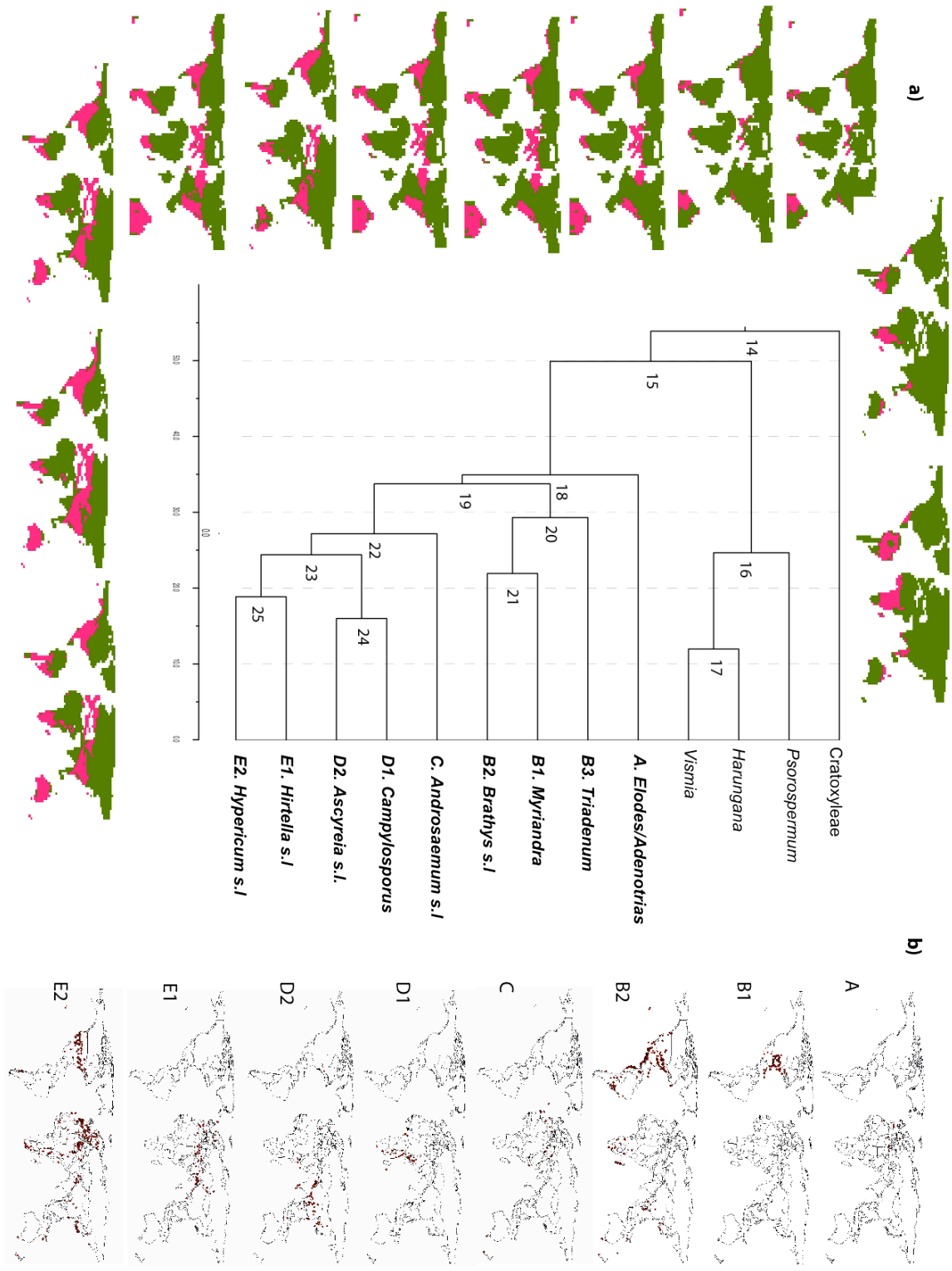


Figure 6. a) Biplot showing the Mahalanobis distance from the optimum niche value inferred for Hypericaceae ancestor (node 14) to the estimated average niche value of all the study phylogenetic nodes (see text). Numbers over circles represent phylogenetic nodes as in Figure 4. Circle size is scaled according with the inferred ancestral tolerances. b) Biplot representing minimum month temperature against minimum month precipitation values as inferred for the *Hypericum* ancestors. Circles represent phylogenetic nodes in subfigure a) and can be identify by the corresponding node number. Circle size is scaled according with the inferred ancestral tolerances. c) Same biplot as in b) but including values of phylogenetic clades in the present.

Figure 7 (next page). a) Ancestral niche reconstruction results and geographic projections. Pruned chronogram of *Hypericum* obtained from Meseguer et al (2013), to represent mayor lineages and outgroup representatives. Pink colour on maps represents the potential distribution inferred for every ancestor in the past. b) Present distribution of major phylogenetic lineages.


a)
Cratoxyleae
Psorospermum
Harungana
Vismia
A. Elodes/Adenotrias
B3. Triadenum
B1. Myriandra
B2. Brathys s.l
C. Androsaemum s.l
D1. Campylosporus
D2. Ascyreia s.l.
E1. Hirtella s.l
E2. Hypericum s.l
14
15
16
17
18
19
20
21
22
23
24
25
50.0
40.0
30.0
20.0
10.0
0.0
b)
A
B1
B2
C
D1
D2
E1
E2

GENERAL DISCUSSION.

GENERAL DISCUSSION.

I. Phylogenetic and taxonomic implications

In Chapter 1 we present a phylogenetic study for the genus *Hypericum* based on different nuclear (ITS) and chloroplast markers (*trnL-trnF*, *trnS-trnG*, *psbA-trnH*) that includes nearly 40% of the nearly 500 species described in the genus. The plastid phylogenies exhibited good resolution and support values, while the nuclear trees were generally less resolved, in agreement with previous studies (Nürk et al., 2013; Crocket et al., 2004; Park & Kim, 2004). We found good levels of congruence between nuclear and plastid signal (i.e., groups that received significant statistical support in the ITS tree were also supported in the cpDNA phylogeny, and viceversa). The main exception was the barcoding marker *psbA-trnH*, which showed some phylogenetic relationships that were not supported by either the nuclear or the other plastid regions. This is explained by the high levels of homoplasy found in *psbA-trnH*, which agree with the complex mutational dynamics described for this plastid region (Borsch & Quand, 2009); nevertheless, additional underlying biological processes such as biparental inheritance and introgression (Greiner et al., 2011) cannot entirely be discarded. We also demonstrating the potential utility of low-copy nuclear genes for solving phylogenetic relationships in *Hypericum* (Chapter 2), especially in cases of incongruence between plastid and ITS signal. The amplification of these nuclear markers had some difficulties, mainly related with primer specificity and the presence of paralogous copies, and we only managed to amplify two of the eight different molecular markers we initially screened across a representative sample of *Hypericum* species. Nevertheless, two of these regions, EMB2567 and PHYC, and especially the latter, are promising in that they produced consistent, well-resolved trees, sometimes increasing the resolution for basal nodes compared with more traditional markers employed in Chapter 1. One potential problem of EMB we found, is that in *Hypericum* it is presented in multiple copies. Therefore, amplification of this region will require further cloning strategies or clade-specific primer design.

In terms of phylogenetic conclusions, this PhD work confirms the division of family Hypericaceae into three well-supported tribes: Vismieae, Cratoxyleae and Hypericeae. It also agrees with previous molecular studies in inferring genus *Hypericum* as non monophyletic, with smaller genera *Triadenum* nested within; *Thornea* appears in a basal polytomy together

with the early divergent sections *Elodes* and *Adenotrias* in the ITS tree. Ruhfel et al. (2011)'s phylogeny, based on different plastid markers, synonymized *Triadenum*, *Thornea and Santomasia* with *Hypericum*, whereas Nürk et al. (2013)'s ITS phylogeny placed *Thornea* as its sister-group. Further molecular markers might be needed to solve generic relationships within Hypericeae, as well as the inclusion of the missing genus *Lianthus*, but if confirmed, these inferences imply major taxonomic rearrangements in which Hypericeae will be a monotypic tribe represented by only genus *Hypericum*, with all other former genera synonymized with it.

Whithin *Hypericum*, our results showed that the species-poor sections *Elodes* and *Adenotrias* are the sister-group to the remaining species in the genus. This result, however, is only weakly supported (pp < 70%) by ITS and plastid markers in Chapter 1, and does not appear in the LCG trees of Chapter 2. It is however supported by the ITS phylogenetic hypothesis of Nürk et al., (2013). What all markers explored here show is a geographic dichotomy between an Old World and a New World clade, forming a clade sister to *Elodes-Adenotrias*. The New World group includes mainly American species from sections *Brathys*, *Myriandra* and *Trigynobrathys*, although some species of this clade have reached the East of Asia, Oceania and Africa. The Old World group is formed by the remaining sections of the genus as section *Adenosepalum*, *Hypericum*, *Hirtella* or *Ascyreia* among others, and has its centre of diversity in the Palearctic.

Another relevant result from the phylogenetic analyses is the lack of congruence between the sectional classification of the genus, based on morphological characters and the molecular results. Of the 36 recognized sections (Robson, 2012), twelve sections were recovered para- or polyphyletic, eight monotypic and only three confirmed to be monophyletic. Among the non-monophyletic sections stand out the species-rich sections *Ascyreia*, *Brathys*, *Hypericum* or *Adenosepalum*. This inconsistency between molecular and morphological evidence points out towards convergent or parallel evolution. We confirmed that this was the case for most of the diagnostic characters analysed in Chapter 1, such as the presence of five stamen fascicles in the androecia or the existence of dark glands. We found that these characters have evolved independently in several clades from three-fasciculate dark-glandless ancestors. Habit form has been used to discriminate the New World sections *Brathys* and *Trigynobrathys* (Robson, 1990). However, our work shows that this trait is very labile in *Hypericum*, and there have been several cases of independent evolution from the shrub to the treelet and herbaceous

forms. These high levels of morphological plasticity presumably allowed *Hypericum* to adapt to different environments, such as the seasonal temperate habitats in the Northern Hemisphere, but it will hinder the construction of a morphologically based classification. Methodological difficulties associated to "long-branch attraction" phenomena or elevated mutational rates could obscure phylogenetic relationships, as shown in Chapters 1 and 2 (e.g., EMB and *psbA-trnH*). In some cases, probabilistic models implemented under incorrect parameter settings may lead to branch length overestimations, especially in multigene analyses when there is high heterogeneity of mutation rates among partitions (Brown et al., 2010). We found that this might also occur in single gene analyses when evolutionary rates significantly differ significantly within gene regions. Modification of the branch length prior "lambda" in the Bayesian analyses of the concatenated cpDNA and ITS datasets helped recover more realistic branch lengths, in accordance with the maximum-likelihood values. Branch length overestimation might not alter phylogenetic relationships, but it dramatically affects analyses based on branch length measurements such as molecular clock dating and diversification tests.

In all, this work stresses the importance of including several unlinked evidences to examine evolutionary relationships, as different molecular markers can produce different topologies, but also as a tool to identify important biological phenomena like convergent evolution both at the molecular and morphological levels.

What comes next? Future phylogenetic studies should focus on identifying new markers to increase resolution both at the basal and within-clade levels. At the basal level, the position of *Thornea* within *Hypericum* should be clarified as well as the placement of sections *Elodes* and *Adenotrias* with respect to the rest of the genus. At the species level, it is necessary to explore the potential of fast evolving molecular markers to increase resolution for some recently radiated phylogenetic clades like the *Hypericum* s.l. and the *Brathys*-group. Topological differences at the species level could also be seen among the study markers. Whether this is the product of methodological artefacts (homoplasy) as suggested here, or the result of biological processes such incomplete lineage sorting, paralogy, or hybridization, remains to be tested. In this sense, new multi-marker coalescent approaches can be appropriate to approximate the true species phylogeny. In addition to the exploration of new molecular markers, increasing the species sampling for poorly represented sections like *Concinna*, *Taeniocarpium*, and *Origanifolium*, and those with large morphological diversity

such as the American sections *Brathys*, *Trigynobrathys*, the Mediterranean *Hirtella*, and the Asian *Ascyreia*, is another priority. Finally, a major aim of any study would be to include the Hypericeae genus *Lianthus* in order to clarify the taxonomic circumscription of Hypericeae and *Hypericum*.

II. *Hypericum* fossil record

Hypericum has a rich fossil record that extends from the Late Eocene to the present. This record comprises some macrofossils (leaves), but mostly consists of microfossil remains, seeds and pollen. Pollen in *Hypericum* is not particularly distinctive, but seeds have a characteristic ornamentation of the testa, ranging from reticulate, scalariform or papillose to smooth, which sometimes characterize entire taxonomical sections (Robson, 1981).

The oldest unequivocal fossils attributed to *Hypericum*, are the remains found in a Late Eocene site in Siberia, assigned to the extinct species *Hypericum antiquum* Balueva & Nikitin (Arbuzova, 2005). *H. antiquum* seeds have a characteristic ribbed (scalariform) sculpting pattern of the testa that is present in extant species of sections *Elodes*, *Brathys*, *Trigynobrathys* and *Drosocarpium*. The next oldest remain belongs to the fossil species *H. septestum*, and was also found in the Eastern Palearctic region in Lower Oligocene deposits (Tomsk, Russia; Arbuzova 2005). This species became a common element in the European fossil record of the genus from the Late Oligocene to the Pliocene. Seeds of *H. septestum* differ from the older *H. antiquum* remains in possessing reticulate patterns in the seed testa. This character state is by far the most abundant within the genus and appears scattered along the phylogeny (Chapter 1).

Although the sculpture pattern of the testa varies across species and clades, we showed in Chapter 1 that this character is not constrained phylogenetically in *Hypericum* and exhibits high levels of homoplasy. With the exception of the oldest fossil *H. antiquum,* this prevents confident assignment of seed remains to phylogenetic clades. Since *H. antiquum* does not exhibit any other particular feature (apomorphy) distinctive from other extant species, we used this fossil to constrain the age of the crown node of *Hypericum*, which is the most recent common ancestor of all sections with a scalariform testa sculpture (*Elodes*, *Brathys*, *Trigynobrathys* and *Drosocarpium*), thus providing a minimum age for the group (Magallón & Sanderson, 2001). This calibration point was used in the dating analyses of Chapter 1, 4 and 5.

The distribution of the fossil record of *Hypericum* is an important tool to gives us clues on its past biology and geographic range. The geographic distribution of paleontological sites with *Hypericum* remains suggest a long history of the genus in the Holarctic region, with presence in the Palearctic by the end of the Eocene, probably as part of the boreotropical forest belt (*H. antiquum*), or of its successor, the mixed-mesophytic forest, as evidenced by the presence of *Hypericum* remains in a Miocene assemblage of this forest in Yunnan (China; Zhao et al., 2004). Fossil remains in other parts of the world, such as America and Africa, do not appear until the Late Cenozoic (Plio-Pleistocene). This is surprising given that many extant species are endemic to those areas. It is also noteworthy that some fossils are assigned to extant species that are not presently distributed in the area where the fossil was found. This could be an error of identification, an evidence of the effect of extinction, or true absence of the genus from those areas. Further paleontological work will help to clarify these issues.

III. Spatio-temporal evolution of *Hypericum*

Along Chapters 1, 3, 4 and 5, we investigated the biogeographic and diversification history of *Hypericum*, as well as the evolution of ecological preferences, using different methodologies and sources of information. The biogeographic history that emerges after the integration of all available sources of evidence is one in which Hypericaceae ancestors originated in Africa in the Early Eocene (ca. 55 Ma), from where they dispersed to Eurasia before 50 Ma (Chapter 4) probably favoured by the tropical climate conditions dominant in the Holarctic at that time (see section on Cenozoic climate change). This period was the warmest in the history of the Cenozoic; at the warmth peak of the Early Eocene Climatic Optimum (EECO) global temperatures were on average 5 to 10 ºC higher than today (Zachos et al., 2008). Entrance into the Holarctic probably occurred along the branch leading to the ancestor of Hypericeae-Vismieae and before the first diversification within *Hypericum* (crown group, see Chapter 1 and 4). From Eurasia, *Hypericum* stem lineages entered into North America before 35 Ma. This dispersal took place probably across Beringia, as revealed by our fossil-based ENM reconstructions (Chapter 4), which show climatically favourable areas in this region over the entire Early Cenozoic. Crown group diversification, the MRCA of all living *Hypericum*, is dated at 35 Ma, coincident with the dramatic drop in temperatures known as the TEE event. This climate cooling produced the decline of the boreotropical forest and the expansion of a more deciduous vegetation, the mixed-mesophytic forest.

Soon after the first diversification of *Hypericum*, a vicariant event isolated the New World and the Old World lineage in North America–Eastern Palearctic and Western Palearctic, respectively, and could have been mediated by the presence of the ancient Turgait Strait separating the western and eastern halves of the Palearctic, as well as the effect of the TEE climate cooling. In the New World group, this vicariance was immediately followed by the extinction of the EP populations before 30 Ma. This could be explained by a dramatic reduction in the percentage of area with favourable climatic conditions in eastern part of Beringia after the Eocene, as inferred from the fossil-based ecological niche reconstructions (Chapter 4). Indeed, the Beringian corridor did not become appropriate again for the dispersal of *Hypericum* lineages until the Pliocene, concurrent with a dramatic increase in global temperatures at the MPW event (Chapter 4). Most of the inferred dispersal and subsequent vicariance events between the Nearcctic and Palearctic lineages are dated around this period, such as *Triadenum*, *H. ascyron*, *H. mutilum* or *H. erectum-H. formosum* (Chapter 1, 4).

The Old World lineage apparently persisted in the Western Palearctic until the beginning of the Miocene, after which several lineages extended their range into the African and Asian continents (Chapter 1). However, the history of *Hypericum* in the Eastern Palearctic (Asian) region is still under debate. There are fossils of *Hypericum* recorded in this region in all geological periods from the Late Eocene to the present (Chapter 3), which suggests a continuous presence of the group in the region. This is not reflected in our biogeographic reconstructions, which either inferred a late colonization (Chapter 1) or an early extinction followed by a late recolonization when the crown-*Hypericum* ancestral area is constrained with the distribution of the oldest fossil in this region (Chapter 4). Unfortunately, with the exception of *H. antiquum*, none of these fossil remains in EP posses specific features for its assignation to particular clades, which precludes their inclusion in biogeographic analyses. Nevertheless, several lines of evidence support the idea that the eastern half of the Palearctic was part of the ancestral area of *Hypericum*, but that the phylogenetic signal has been lost as the result of extinction. Fossil-based ENM reconstructions (Chapter 4) show that the percentage of climatic favourable area in this region was dramatically reduced from the Early Miocene to the present, compared with the Western Palearctic and the Nearctic regions in which the amount of favourable area remained relatively constant through time. The only exception is the southeastern region of EP, which seems to have been favourable throughout *Hypericum* history. Trait-dependent diversification

models also indicated that lineages living in the Eastern Palearctic region presented higher speciation and extinction rates than those living in other areas, suggesting a higher species turnover in this region (Chapter 5). Although more investigation is needed, under this interpretation, the *Ascyreia* s.l. lineage, now mainly distributed in South East Asia and the Himalayan mountain range, might represent a relict group that found refuge in small portions of Eastern Asia during the Late Tertiary climatic fluctuations.

Our biogeographic results indicate that the entrance of *Hypericum* into the southern parts of the Southern Hemisphere did not occur before the Late Miocene and probably through the newly uplifted mountain ranges in South America and Eastern Africa (Chapters 1, 4). Our fossil-based ENM models suggest that the southernmost temperate regions of these continents and Australia presented favourable conditions for *Hypericum* throughout its evolutionary history (Chapter 4). This stands in contrast with the current low number of endemic species in South America and Australia, which harbour mostly species widespread with other regions. Although these former Gondwanan regions became geographically connected to the Holarctic in the Late Tertiary - through the Sunda region and the Panama Isthmus, respectively - the tropical equatorial belt in South America and South East Asia probably acted as a climatic barrier, preventing *Hypericum* lineages to colonize the southernmost temperate regions. Africa constitutes an exception to this pattern because it harbours more endemic species and higher clade diversity than South America or Australia. One explanation is that the aridification trend that affected this continent from the Mid-Tertiary onwards, proceeding from south to north, gradually led to the replacement of ancestral tropical forests by woodlands and savannah (Senut et al., 2009). This provided a new ecological corridor across North Africa in the Miocene, which, together with the collision of Arabia (16 Ma ago), allowed Eurasian Old World lineages to migrate into Africa (e.g., the *Androsaemum* group). This connection was broken in the Pliocene, when an increase in aridification led to the formation of large deserts across North Africa and Arabia. However, mountain building in Eastern Africa during the Pliocene (Sepulchre et al., 2006) provided the Afromontane and *Hypericum* groups with new cool-temperate habitats to disperse southwards and diversify.

IV. Drivers of diversification: ecological innovation versus ancestral resilience

The results presented in this thesis have implications beyond the specific questions addressed for *Hypericum*, providing information of the response of biodiversity to environmental pressures (climate and geological change) in evolutionary time scales.

The extent to which species are labile to preserve or to evolve ecological is an intensely debated subject (Wiens & Donoghue, 2004, Donoghue 2008, Wiens et al., 2010, Wiens & Graham, 2005, Peterson et al., 1999, Peterson, 2011, Prinzing et al., 2001, Crisp et al., 2009), which has implications on the ability of the species to adapt to new environments. Ecologists propose that the answer to this question is scale dependent (Wiens et al., 2010). Niche conservatism might prevail at small temporal scales, i.e., sister species comparisons show significant niche preservation between species pairs (Peterson et al., 1999). While niche differences are expected to accumulate through time, and to be observed within higher taxonomic ranks (e.g. families, orders; Peterson et al., 1999; Peterson 2011; Cronquist, 1968). However, phylogeneticists consider that broad niche preferences can be conserved over very long periods of time (Wiens&Graham, 2005; Wiens & Donoghue, 2004; Donoghue 2008; Wiens et al., 2010; Crisp et al., 2009). This is based on the observation that ancient clades are generally confined to one biome and absent from other biotic regions, as major angiosperm clades are restricted to either temperate or tropical forest, grasslands or desserts (Donoghue, 2008; Wiens & Graham, 2005). Theoretical predictions justify that niche conservatism over long time scales might be a consequence of rates of adaptation to new environments (rates of trait evolution) evolving generally slower than rates of extinction (Holt & Gaines, 1992). However, at large scales, only a few studies have really investigated this question analytically (Crisp et al., 2009), and always basing the reconstruction of ancestral preferences on extant evidences. One problem of this approach, as commented in the introduction, is that it suffers from uniformitarianism assumptions (Hutton 1794; Gould 1965; Lyell, 1830). Given that the probability of estimating a particular parameter for the ancestors is conditional to the present value of the descendants, ancestral niche inferences could not been done without assuming, to a certain extent, that niches are conserved through time. In this thesis, we overcome previous limitations using a fossil-based ENM approach to showcase that niche climatic preferences can be conserved over evolutionary time scales of tens of millions of years (Chapter 4). Present tolerances of *Hypericum* were already achieved by the end of the Eocene, and were maintained relatively stable along its 35 Ma of evolutionary history since the first

reconstructed diversification event. This does not exclude that niche evolution occurred earlier in the history of the group. In fact, a change in climatic preferences seem to have taken place along the branch leading from *Hypericum* stem lineages (the divergence from its sister-group, Vismieae) to crown-node *Hypericum,* as inferred by ancestral niche reconstruction based on extant occurrences (Chapter 5). Interestingly, preferences of the *Hypericum* stem ancestors were reconstructed as intermediate between descendant clades with tropical and temperate affinities: they seem to have lived under lower temperature and higher precipitation regimes than tropical taxa. This could reflect the preferences of a boreotropical group and is in accordance with what we know of this forest, a mixture of hardwood deciduous and evergreen taxa with no analogy in the present.

Although the pooled niche of the genus has apparently been conserved from the Eocene to the present, other episodes of niche evolution occurred within *Hypericum*, with major phylogenetic clades specializing along different subsets of the ancestral generic niche (Chapter 4, 5). There is a clear, observable tendency towards increased cold tolerance along the evolution of *Hypericum*, from the Early Eocene boreotropical Hypericaceae ancestors, living under an average minimum monthly temperature of 10ºC, to the cold-temperate Hypericum crown lineages at 4ºC, to present clades such as *Hypericum* s.l or *Hirtella* s.l, which present minimum monthly temperatures under 0ºC. There is also a weaker trend towards increasing drought resistance along the lineages of *Hypericum* and Vismieae relative to their ancestors. This agrees well with the global cooling trend that started by the Mid Cenozoic, promoting the appearance of cooler and more seasonal environments in the Holarctic (Zachos et al., 2008) and a trend towards increased aridification in Africa (Plana 2004).

However, besides these general trends, the evolution of ecological preferences has not been directional in *Hypericum*, but heterogeneous across phylogenetic clades, with departures and approaches to ancestral conditions along different dimensions of the climatic niche. Some groups maintained similar climatic temperature preferences to the ancestors (*Elodes*/*Adenotrias*, *Myriandra*), while others became adapted to new temperature and precipitation regimes in the subtropical mountains, exploring different dimensions of the ecological space. For example, the Afromontane *Campylosporus* evolved higher drought tolerance but increased its minimum monthly temperature towards more tropical conditions (similar to Vismieae), while Asian (Himalayan) *Ascyreia* its tolerance to cold temperatures

but showed lower tolerance to drought. A similar pattern can be found between Nearctic *Myriandra* and the Andean *Brathys*, with the latter living in wetter environments more similar to the Hypericaceae ancestors. This heterogeneity or lack of "directionality" in the exploration of the niche space, with bounces back and forth from the intermediate ancestral conditions, can be found also between the ancestral nodes

These events of niche divergence, some of them concurrent with important climatic events (TEE) led some authors to associate ecological innovation with the rapid diversification of genus *Hypericum* (Meseguer et al., 2013; Nurk et al., 2013). The key-innovation hypothesis (KID) suggests that a change in climatic tolerances to deal with the new temperature regimes in the Holarctic or the appearance of the herbaceous habit form were "adaptive breakthroughs" that allowed the genus to diversify rapidly and ultimately being responsible for its evolutionary success. Specifically, ecological differentiation have been associated with diversification (Graham et al., 2004; Moritz et al., 2000; Kozak & Wiens, 2007), but only in some particular cases the correlation between the rate at which the niche diverges and the rate at which species accumulates through time has been demonstrated (Schnitzler et al., 2012; Kozak & Wiens, 2010). In *Hypericum*, we found that neither ecological innovation nor a change in habit form promoted diversification (Peterson et al., 1999; Warren et al., 2008), although it may have contributed to decrease the extinction risk under scenarios of climatic instability (Chapter 5). The present diversity of *Hypericum* is not particularly high given its age and the background diversification rate of the clusioid clade. In agreement with the "time-to-speciation effect" (TSE) hypothesis, the steady accumulation of lineages during the last 35 Ma explains better the large species richness in this group than events of rapid diversification. However, even if *Hypericum* did not diversify rapidly, it is still richer than other clusioid genera of the same old-age (e.g., Wedellina).

Why then *Hypericum* is so species rich compared to these relatives? The answer seems to lie in its high "resilience", i.e., the ability to cope with disturbance and stress without showing negative effects, which might imply bouncing back to previous states, undergo change while still retaining the same function and capacity, or the coexistence of multiple stable states or regimes (Holling, 1973). *Hypericum* evolved from tropical ancestors in Gondwana and survived in the Holarctic through all drastic climate changes of the Cenozoic which extirpated most of the boreotropical and temperate taxa from the area. This resilience is probably rooted in the high genetic plasticity of the ancestral boreotropical lineages, which

presented broad niche tolerances. *Hypericum* has diversified steadily during the last 35 million years, without sharp increases or decreases of diversification, in some cases coping with the loss of diversity by geographic movement to new areas, in others by "bouncing back" to previous ancestral states, and ultimately evolving new traits. This plasticity seems to have been preserved over time: present clades have the same capacity to explore very different dimensions of the niche space than their ancestors, which has important implications for the ability of *Hypericum* to cope with future climate change.

Much interest has been focused on the study of adaptive radiations and key innovations to understand the generation of biodiversity. These studies usually deal with young, recently diverged "sprinter" groups. However, results from this work demonstrate that ancient lineages like *Hypericum*, which have survived through the climate changes of the Cenozoic, also merit attention for what they can tell us about the role of "resilience" and "plasticity" in coping with extinction. Like a *marathon runner*, keeping the pace during a long-distance running event, *Hypericum* has been able to survive and diversify through all major climatic changes of the Cenozoic, which has allowed it its current large diversity and cosmopolitan geographic distribution.

Finally, the results presented in this thesis have implications beyond the specific questions addressed for *Hypericum*, providing information of the response of biodiversity to environmental pressures (climate and geological change) in evolutionary time scales. Specifically, using a fossil based approach we showcase that niche preferences can be conserved over evolutionary time scales of millions of years, which has implications into the niche conservatism debate: to which extent species are labile to preserve or to evolve ecological preferences? (Donoghue, 2008). We also aimed at stressing the importance of an integrative approach in evolution, which by incorporating all present and extinct sources of evidence (ecological, paleontological, and phylogenetic), may improve our understanding of the history of organisms. The use of the fossil record and ecological preferences to inform biogeographic and evolutionary reconstructions may help to overcome the bias introduced by the assumption of "actualism" in evolutionary biology. This new integrative approach might be applicable to other organisms to reconstruct the spatio-temporal evolution of extant and extinct lineages through time.

REFERENCE LIST

REFERENCE LIST

Ackerly DD (2004) Adaptation, niche conservatism, and convergence: Comparative studies of leaf evolution in the California chaparral. Am Nat 163:654–671.

Arbuzova O (2005) *Hypericum* L. In: Budantsev, L. (eds.), Iskopaemye tsvetkovye rastenija Rossii i sopredel'nyh gosudarstv [Fossil flowering plants of Russia and adjacent countries], Vol. 4 Nyctaginaceae-Salicaceae. Izdatelstvo Nauka Leningradskoe otd-nie, 1974-. Leningrad. (In Russian).

Barnosky AD (2001) Distinguishing the effects of the Red Queen and Court Jester on Miocene mammal evolution in the northern Rocky Mountains. Journal of Vertebrate Paleontology 21:172–185.

Benton MJ (2009) The Red Queen and the Court Jester: species diversity and the role of biotic and abiotic factors through time. Science 323: 728–732.

Borsch T, Quandt D (2009) Mutational dynamics and phylogenetic utility of noncoding chloroplast DNA. Plant Systematics and Evolution 282: 169–199.

Broennimann O, Guisan A (2008) Predicting current and future biological invasions: both native and invaded ranges matter. Biology Letters 4: 585–589.

Brown JM, Hedtke SM, Lemmon AR, Lemmon EM (2010) When trees grow too long: investigating the causes of highly inaccurate Bayesian branch-length estimates. Systematic Biology 59: 145–161.

Buerki S et al. (2011) An evaluation of new parsimony-based versus parametric inference methods in biogeography: a case study using the globally distributed plant family Sapindaceae. J Biogeogr 38: 531–550.

Carine MA, Christenhusz MJM (2010) About this volume: the monograph of *Hypericum* by Norman Robson. Phytotaxa 4: 1–4.

Choisy JD (1821) Prodromus d'une monographic de la famille des Hypericacees, Geneva.

Crane PR, Lidgard S (1990). Angiosperm diversification and paleolatitudinal gradients in Cretaceous floristic diversity. Science 246: 675–678.

Crepet WL, Nixon KC (1998) Fossil Clusiaceae from the late Cretaceous (Turonian) of New Jersey and implications regarding the history of bee pollination. American Journal of Botany 85: 1122–1133.

Crisp M, Arroyo MTK, Cook LG, Gandolfo MA, Jordan GJ, McGlone MS, Weston PH, Westoby M, Wilf P, Linder HP (2009) Phylogenetic biome conservatism on a global scale. Nature 458: 754-756.

Crockett SL, Douglas AW, Scheffler BE, Khan IA (2004) Genetic profiling of Hypericum (St. John's Wort) species by nuclear ribosomal ITS sequence analysis. Planta Medica 70: 929–935.

Cronquist A (1968) The Evolution and Classification of Flowering Plants (Houghton Mifflin, Boston)

Darwin CR (1845) Journal of researches into the natural history and geology of the countries visited during the voyage of H.M.S. Beagle round the world, under the Command of Capt. Fitz Roy, R.N. 2d edition. London: John Murray.

Davis CC, Webb CO, Wurdack KJ, Jaramillo CA, Donoghue MJ (2005) Explosive radiation of malpighiales supports a mid-Cretaceous origin of modern tropical rain forests. American Naturalist 165: E36–E65.

Donoghue MJ, Smith SA (2004) Patterns in the assembly of temperate forests around the Northern Hemisphere. Philos Trans R Soc Lond B Biological Sciences 359:1633–1644.

Donoghue MJ (2008) A phylogenetic perspective on the distribution of plant diversity. Proc Natl Acad Sci USA 105:11549-11555.

Doyle JJ (1992) Gene trees and species trees: molecular systematics as one character taxonomy. Systematic Botany 17: 144–163.

Engler A (1925) Guttiferae. In: Engler A & Prantl K (Eds.) Die natürlichen Pflanzenfamilien, ed. 2, 21. Engelmann, Leipzig, pp. 154–237.

Ezard THG, Aze T, Pearson PN, Purvis A (2011) Interplay between changing climate and species ecology drives macroevolutionary dynamics. Science 332: 349–351.

FitzJohn RG, Maddison WP, Otto SP (2009) Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. Syst Biol 58: 595–611.

Forster JR (1778) Observations made during a voyage round the World, in Physical Geography, Natural History, and Ethnic phylosophy. London: G. Robinson.

Goldberg EE, Lancaster LT, Ree RH (2011). Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. Syst Biol 60: 451–465.

Gould SJ (1965) Is uniformitarianism necessary? American Journal of Science 263: 223–228.

Graham CH, Ron SR, Santos JC, Schneider CJ, Moritz C (2004) Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. Evolution. 58: 1781–1793.

Greiner S, Rauwolf U, Meurer J, Herrmann RG (2011) The role of plastids in plant speciation. Molecular Ecology 20: 671–691.

Heenan PB (2008) Three newly recognised species of Hypericum (Clusiaceae) from New Zealand. New Zealand Journal of Botany 46: 547-558.

Herrera CM (1992) Historical effects and sorting processes as explanations for contemporary ecological patterns: Character syndromes in Mediterranean woody plants. Am Nat 140: 421–446.

Hillis DM (1995) Approaches for assessing phylogenetic accuracy. Syst Biol 44: 3–16.

Holling CS (1973) Resilience and stability of ecological systems. Annual Review of Ecology and Systematics 4: 1–2.

Holt RD, Gaines MS (1992). Analysis of adaptation in heterogeneous landscapes: implications for the evolution of fundamental niches. Evol Ecol 6: 433–447.

Hubbell SP (2001) The Unified Neutral Theory of Biodiversity and Biogeography. Princeton Univ Press, Princeton.

Huelsenbeck JP, Bollback JP (2001) Empirical and hierarchical Bayesian estimation of ancestral states. Systematic Biology 50: 351–366.

Hughes C, Eastwood R (2006) Island radiation on a continental scale: exception rates of plant diversification after uplift of the Andes. Proc Natl Acad Sci USA 103: 10334–10339.

Humboldt A (1820) Voyage aux regions equinoxiales du Nouveau Continent. In N. Mazé (Eds.). Paris.

Hutchinson GE (1957). Concluding remarks. Cold Springs Harbor Symposia on Quantitative Biology 22: 415–427.

Hutton (1794) An investigation of the principles of knowledge and of the progress of reason, from sense to science and philosophy. Strahan & Cadell, Edinburg.

Jaubert C, Spach E (1842) Illustrationes Plantarum orientalium, I. Paris.

Judd WS, Sanders RW, Donoghue MJ (1994) Angiosperm family pairs: preliminary phylogenetic analyses. Harvard Papers in Botany 5: 1–51.

Keller R (1925) Hypericum. In: Engler A, Prantl K (Eds.). Die natürlichen Pflanzenfamilien. Engelmann, Leipzig, pp. 175–183.

Kimura Y (1951) Hypericaceae. In: Nakai T, Honda M (Eds.). Nova Flora Japonica, Tokyo, Sanseido.

Kozak KH, Wiens JJ (2010) Accelerated rates of climatic-niche evolution underlie rapid species diversification. Ecology Letters 13: 1378–1389.

Lemey P, Rambaut A, Drummond AJ, Suchard MA (2009) Bayesian phylogeography finds its roots. PLoS Computational Biology 5.

Lieberman BS (2002) Phylogenetic biogeography with and without the fossil record: gauging the effects of extinction and paleontological incompleteness. Palaeogeography, Palaeoclimatology, Palaeoecology 2726: 1–14.

Lieberman BS (2003) Unifying theory and methodology in biogeography. Evol Biol 33: 1–25.

Linder & Sanmartín, in prep

Lomolino MV, Riddle BR, Brown JH (2005) Biogeography. Sinauer Associates, Sunderland, MA.

Lyell C (1830-1833). Principles of Geology. J. Murray, London.

Maddison WP, Midford PE, Otto SP (2007) Estimating a binary character's effect on speciation and extinction. Syst Biol 56: 701–710.

Magallon S, Sanderson MJ (2001) Absolute diversification rates in angiosperm clades. Evolution 55: 1762–1780.

Maguire KC, Stigall AL (2009) Using ecological niche modelling for quantitative biogeographic analysis: a case study of Miocene and Pliocene Equinae in the Great Plains. Paleobiology 35: 587–611.

Mao K, et al. (2012) The distribution of living Cupressaceae reflects the breakup of Pangea. Proc Natl Acad Sci USA 109: 7793–7798.

Matzk F, Meister A, Brutovska R, Schubert I (2001) Reconstruction of reproductive diversity in *Hypericum perforatum* L. opens novel strategies to manage apomixis. The Plant Journal 26: 275–282.

Moritz C, Patton JL, Schneider CJ, Smith TB (2000). Diversification of rainforest faunas: an integrated molecular approach. Annu Rev Ecol Syst 31: 533–563.

Morlon H, Parsons TL, Plotkin J (2011) Reconciling molecular phylogenies with the fossil record. Proc Natl Acad Sci USA 108: 16327–16332.

Nauheimer L, Metzler D, Renner SS (2012) Global history of the ancient monocot family Araceae inferred with models accounting for past continental positions and previous ranges based on fossils. New Phytol 195: 938–950.

Nee S, Holmes EC, May RM, Harvey PH (1994) Extinctions rates can be estimated from molecular phylogenies. Philos Trans R Soc Lond 344: 77–82.

Nee S, May RM, Harvey PH (1994) The reconstructed evolutionary process. Philos Trans R Soc Lond B Biol Sci 344: 305–311.

Nogues-Bravo D, Rodriguez J, Hortal J, Batra P, Araujo MB (2008) Climate change, humans, and the extinction of the woolly mammoth. PLoS Biol 6:e79.

Nürk NM, Blattner FR (2010) Cladistic analysis of morphological characters in *Hypericum* (Hypericaceae). Taxon 59: 1495–1507.

Nürk NM, Madriñán S, Carine MA, Chase MW, Blattner FR (2013) Molecular phylogenetics and morphological evolution of St. John's wort (*Hypericum*; Hypericaceae). Molecular Phylogenetics and Evolution 66: 1-16.

Park S, Kim K (2004) Molecular phylogeny of the genus *Hypericum* (Hypericaceae) from Korea and Japan: evidence from nuclear rDNA ITS sequence data. Journal of Plant Biology 47: 366–374.

Peterson AT (2011) Ecological niche conservatism: a time structured review of evidence. J Biogeogr 28: 817–27.

Peterson AT, Soberón J, Sánchez-Cordero V (1999) Conservatism of ecological niches in evolutionary time. Science 285: 1265–1267.

Pilepić KH, Balić M, Blažina N (2011) Estimation of phylogenetic relationships among some Hypericum (Hypericaceae) species using internal transcribed spacer sequences. Plant Biosystems 145: 81–87.

Prinzing A, Durka W, Klotz S (2001). The niche of higher plants: evidence for phylogenetic conservatism. Proc R Soc Lond B 268: 2383–2389.

Qian, H. & Ricklefs, R.E. (2004) Geographical distribution and ecological conservatism of disjunct genera of vascular plants in eastern Asia and eastern North America. Journal of Ecology, 92, 253–265.

Rabosky DL (2006) Likelihood methods for detecting temporal shifts in diversification rates. Evolution 60: 1152–1164.

Rabosky DL (2009) Ecological limits and diversifi cation rate: Alternative paradigms to explain the variation in species richness among clades and regions. Ecology Letters 12: 735 – 743 .

Ramos-Nuñez ÁF (1993) *Hypericum* L. In: Castroviejo S et al., (Eds.). Flora iberica 3: 157–185. Real Jardín Botánico, CSIC, Madrid.

Raxworthy CJ, Ingram CM, Rabibisoa N, Pearson RG (2007) Applications of ecological niche modeling for species delimitation: a review and empirical evaluation using day geckos (Phelsuma) from Madagascar. Syst Biol 56: 907–923.

Ree RH, Smith SA (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst Biol 57: 4–14.

Ree RH, Moore BR, Webb CO, Donoghue MJ (2005) A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. Evolution 59: 2299–2311.

Ricklefs RE (2006) Evolutionary diversification and the origin of the diversity-environment relationship. Ecology 87(7) Supplement: S3–S13.

Robson NKB (1977) Studies in the genus *Hypericum* L. (Guttiferae). 1. Infrageneric classification. Bulletin of the British Museum (Natural History), Botany Series 5: 295–355.

Robson NKB (1981) Studies in the genus *Hypericum* L. (Guttiferae). 2. Characters of the genus. Bulletin of the British Museum (Natural History), Botany Series 8: 55–226.

Robson NKB (1985) Studies in the genus *Hypericum* L. (Guttiferae). 3. Sections 1. *Campylosporus* to 6a. *Umbraculoides*. Bulletin of the British Museum (Natural History), Botany Series 12: 163–211.

Robson NKB (1987) Studies in the genus *Hypericum* L. (Guttiferae). 7. Section 29. *Brathys* (part 1). Bulletin of the British Museum (Natural History), Botany Series 16: 1–106.

Robson NKB (1990) Studies in the genus *Hypericum* L. (Guttiferae). 8. Sections 29. *Brathys* (part 2) and 30. *Trigynobrathys*. Bulletin of the British Museum (Natural History), Botany Series 20: 1–151.

Robson NKB (1996) Studies in the genus *Hypericum* L. (Guttiferae). 6. Sections 20. *Myriandra* to 28. *Elodes*. Bulletin of the British Museum (Natural History), Botany Series 26: 75–217.

Robson NKB (2001) Studies in the genus *Hypericum* L. (Guttiferae). 4(1). Sections 7. *Roscyna* to 9. *Hypericum* sensu lato (part 1). Bulletin of the British Museum (Natural History), Botany Series 31: 37–88.

Robson NKB (2002) Studies in the genus *Hypericum* L. (Guttiferae). 4(2). Section 9. *Hypericum* sensu lato (part 2): subsection 1. *Hypericum* series 1, *Hypericum*. Bulletin of the Natural History Museum, London (Botany) 32: 61–123.

Robson NKB (2006) Studies in the genus *Hypericum* L. (Clusiaceae). Section 9. *Hypericum* sensu lato (part 3): subsection 1. *Hypericum* series 2. *Senanensia*, subsection 2. *Erecta* and section 9b *Graveolentia*. Systematics and Biodiversity 4: 19–98.

Robson NKB (2010a) Studies in the genus *Hypericum* L. (Hypericaceae). 5(1). Sections 10. *Olympia* to 15/16. *Crossophyllum*. Phytotaxa 4: 5–126.

Robson NKB (2010b) Studies in the genus *Hypericum* L. (Hypericaceae). 5(2). Section 17. *Hirtella* to 19. *Coridium*. Phytotaxa 4: 127–258.

Robson NKB (2012) Studies in the genus *Hypericum* L. (Hypericaceae) 9. Addenda, corrigenda, keys, lists and general discussion. Phytotaxa 72: 1–111.

Ronquist F (2004) Bayesian inference of character evolution. Trends in Ecology and Evolution 19: 475–481.

Ronquist F, Sanmartín I (2011) Phylogenetic methods in historical biogeography. Ann Rev Ecol Evol Syst 42: 441–464

Ruhfel BR (2011) Systematics and Biogeography of the Clusioid Clade (Malpighiales). Harvard University, Cambridge, Massachusetts.

Ruhfe BR, Bittrich V, Bove CP, Gustafsson MHG, Philbrick CT, Rutishauser R, et al. (2011) Phylogeny of the clusioid clade (Malpighiales): evidence from the plastid and mitochondrial genomes. American Journal of Botany 98: 306– 325.

Salvo G, Ho SYW, Rosenbaum G, Ree R, Conti E (2010) Tracing the temporal and spatial origins of island endemics in the Mediterranean region: A case study from the citrus family (Ruta L., Rutaceae). Systematic Biology 59: 705–722.

Sang T (2002) Utility of Low-Copy Nuclear Gene Sequences in Plant Phylogenetics. Critical Reviews in Biochemistry and Molecular Biology 37: 121–147

Sanmartín I, Enghoff H, Ronquist F (2001) Patterns of animal dispersal, vicariance and diversification in the Holarctic. Biological Journal of the Linnean Society 73: 345–390.

Sanmartín I, van der Mark P, Ronquist F (2008) Inferring dispersal: a Bayesian approach to phylogeny-based island biogeography, with special reference to the Canary Islands. Journal of Biogeography 35: 428–449.

Sanmartín I (2012) Historical Biogeography: Evolution in Time and Space. Evo Edu Outreach 5: 555–568. DOI 10.1007/s12052-012-0421-2.

Sanmartín I, Anderson CL, Alarcon M, Ronquist F, Aldasoro JJ (2010) Bayesian island biogeography in a continental setting: the Rand Flora case. Biology Letters 6: 703–707.

Schnitzler J, Graham CH, Dormann CF,Schi_ers K, Linder HP (2012) Climatic niche evolution and species diversi_cation in the Cape flora, South Africa. J Biogeogr 39: 2201–2211.

Schnitzler J, Graham CH, Dormann CF, Schiffers K, Linder HP (2012) Climatic niche evolution and species diversi_cation in the Cape flora, South Africa. J Biogeogr 39: 2201–2211.

Senut B, Pickford M, Sägalen L (2009) Neogene desertification of Africa. Compt Rend Geosci 341: 591–602.

Sepulchre P, Ramstein G, Fluteau F, Schuster M, Tiercelin JJ, Brunet M (2006) Tectonic uplift and Eastern Africa aridification. Science 313: 1419–1423.

Silvestro D, Schnitzler J, Zizka G. (2011) A Bayesian framework to estimate diversification rates and their variation through time and space. BMC Evol Biol 11: 311.

Smith SA, Donoghue MJ (2010) Combining Historical Biogeography with Niche Modeling in the Caprifolium Clade of Lonicera (Caprifoliaceae, Dipsacales). Syst Biol 59: 1–20.

Spach E (1836a). Conspectus monographiae Hypericacearum. Annales des Sciences Naturelles – Botanique et Biologie Vegetale 5: 349–369.

Spach E (1836b). Hypericacearum monographiae fragmenta. Annales des Sciences Naturelles – Botanique et Biologie Vegetale 5: 156–176.

Stadler T (2011). Mammalian phylogeny reveals recent diversification rate shifts. Proc Natl Acad Sci USA 108: 6187–6192.

Stevens PF (2001–onwards) Angiosperm Phylogeny Version 2009, Website: http://www.mobot.org/MOBOT/research/APweb/

Stevens PF (2007) Hypericaceae. In: Kubitzki K (Ed.), The Families and Genera of Vascular Plants. Springer, Berlin, Heidelberg, pp. 194–201.

Stigall AL (2012) Using ecological niche modelling to evaluate niche stability in deep time. J Biogeogr 39: 772–781.

Tiffney BH (1985a) Perspectives on the origin of the floristic similarity between eastern Asia and eastern North America. Journal of the Arnold Arboretum 66: 73–94.

Tiffney BH (1985b) The Eocene North Atlantic land bridge: its importance in Tertiary and modern phytogeography of the Northern Hemisphere. Journal of the Arnold Arboretum 66: 243–273.

van Valen L (1973) A new evolutionary law. Evol Theory 1: 1-30.

Wallace AR (1852) On the monkeys of the Amazon. Proceedings of the Zoological Society of London 20: 107–110.

Warren DL, Glor RE, Turelli M (2008). Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution 62: 2868–2883.

Weaver KF, Anderson T, Guralnick R (2006) Combining phylogenetic and ecological niche modelling approaches to determine distribution and historical biogeography of Black Hills mountain snails (Oreohelicidae). Divers Distrib 12(6): 756–66.

Wen J (1999) Evolution of eastern Asian and eastern North American disjunct distributions in flowering plants. Annual Review of Ecology and Systematics 30: 421–455

Wiens JJ, Graham CH (2005) Niche conservatism: Integrating evolution, ecology, and conservation biology. Annu Rev Ecol Evol Syst 36: 519–539.

Wiens JJ (2007) Global patterns of diversification and species richness in amphibians. Am Nat 170: S86–S106.

Wiens JJ, Ackerly DD, Allen AP, Anacker BL, Buckley LB, Cornell HV, Damschen EI, Davies TJ, Grytnes JA, Harrison SP, Hawkins BA, Holt RD, McCain CM, Stephens PR (2010) Niche conservatism as an emerging principle in ecology and conservation biology. Ecology Letters 13: 1310–1324.

Wiens JJ, Donoghue MJ (2004) Historical biogeography, ecology and species richness. Trends Ecol Evol 19: 639–644.

Willis KJ, MacDonald GM (2011) Long-term ecological records and their relevance to climate change predictions for a warmer world. Annu Rev Ecol Evol Syst 42: 267–87.

Wolfe JA (1975) Some aspects of plant geography of the northern hemisphere during the Late cretaceous and Tertiary. Annals of the Missouri Botanical Garden 62: 264–279.

Xiang QY, Soltis DE, Soltis PS (1998) The eastern Asian and eastern and western north American floristic disjunction: congruent phylogenetic patterns in seven diverse genera. Molecular Phylogenetics and Evolution 10: 178–190.

Yesson C, Culham A (2006) Phyloclimatic modeling: combining phylogenetic and bioclimatic modeling. Syst Biol 55: 785–802.

Zachos JC, Dickens GR, Zeebe RE (2008) An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. Nature 451: 279–283.

Zhao LC, Wanga YF, Liua CJ, Li C (2004) Climatic implications of fruit and seed assemblage from Miocene of Yunnan, southwestern China. Quaternary International 117: 81-89.

CONCLUSIONES GENERALES

CONCLUSIONES GENERALES

El objetivo principal de esta Tesis es investigar los patrones de distribución y riqueza de especies, poniendo un énfasis especial en el efecto que tienen el clima y el cambio geológico sobre los procesos que modelan la diversidad. Para este objetivo estudiamos el género *Hypericum* en tanto que organismo modelo. *Hypericum* es un gran género que tiene una distribución mundial, pero lo que hace más interesante al grupo es que *Hypericum* presumiblemente evolucionó desde unos antepasados tropicales provenientes de Gondwanna y sobrevivió en el Holartico a pesar de todos los cambios climáticos que ocurrieron durante el Cenozoic. La pregunta clave de esta tesis es la siguiente: ¿cómo asimiló *Hypericum* los acontecimientos geológicos y climáticos acaecidos en el pasado hasta llegar a conformar su diversidad y su distribución actuales? Para responder a esta pregunta hemos realizado inferencias del pasado utilizando evidencias del presente junto a material fósil.

A continuación incluimos las conclusiones más importantes de este trabajo:

1. La familia Hypericaceae está bien delimitada e incluye tres tribus: Vismieae, Cratoxyleae and Hypericeae (a la cual pertenece *Hypericum*)
2. *Hypericum* no es monofilético, incluyendo el género *Triadenum* de la tribu Hypericeae. Sin embargo, nuestros resultados no son concluyentes en relación a la posición de *Thornea.*
3. Dentro de *Hypericum* las secciones pobres en especies *Elodea* y *Adenotrias* forman el grupo hermano de una dicotomía geográfica entre especies del Nuevo y del Viejo Mundo.
4. El linaje del Nuevo Mundo incluye principalmente especies Americanas de las secciones *Brathys*, *Myriandra* y *Trigynobrathys*, aunque algunas especies de este clado han alcanzado Asia Oriental, Oceanía y África.
5. El linaje del Viejo Mundo está formado por las secciones restantes del género y su centro de diversidad se sitúa en el Paleártico.
6. La clasificación tradicional de *Hypericum* en 36 secciones morfológicas no se mantiene en base a los resultados moleculares, donde encontramos que muchas de las secciones estudiadas son para- o polyphyleticas. La inconsistencia entre caracteres morfológicos y moleculares se debe probablemente a la evolución convergente en rasgos morfológicos.

7. La estimación de la longitudes de la rama es problemática en análisis Bayesianos cuando existe una elevada heterogeneidad en las tasas de mutación de los diferentes marcadores moleculares. También hemos descubierto que la longitud de las ramas tiende a sobre-estimarse cuando las tasas de evolución difieren significativamente dentro un mismo marcador molecular. La modificación del prior "lambda", que controla las longitudes de la rama, puede ayudar a calcular longitudes de rama más realistas equivalentes a las obtenidas mediantes análisis de Máxima Verosimilitud.
8. Los genes nucleares PHYC y EMB6572 son prometedores para realizar reconstrucciones filogenéticas en *Hypericum*, ya que producen filogenias consistentes, con niveles de resolución y soporte comparables o superiores a otros marcadores nucleares y plastidiales clásicos.
9. *Hypericum* tiene un registro fósil rico que se extiende desde el Eoceno tardío hasta el presente. Este registro comprende algunos macrofósiles (hojas), aunque en su mayor parte consiste de restos microfósiles, semillas y polen.
10. Los registros fósiles más antiguos que pueden atribuirse al género *Hypericum* son del Eoceno Superior (40.4 – 33.9) ycorresponden a las semillas encontradas en Siberia Occidental y descritas como la especie extinta *Hypericum antiquum* Balueva & Nikitin.
11. La edad y distribución de los yacimientos con restos fósiles de *Hypericum* nos permite aproximar la paleodistribución del grupo, mostrando que *Hypericum* se originó en el este del Paleártico, durante el Eoceno, y que posiblemente formaba parte bien del bosque boreotropical o de su sucesor, el bosque mixto-mesofítico, que cubrió el Holártico desde el Paleoceno hasta el Mioceno. Sin embargo, no aparecen restos fósiles en otros continentes, tales como América y África, hasta el final del Cenozoico. Esto puede ser producido por el sesgo del registro fósil, o indicar una verdadera ausencia del género en estas áreas.
12. Los análisis biogeográficos muestran que los orígenes de la familia Hypericaceae se remontan al Eoceno temprano en África, (ca. 55 Ma). Desde África, estos linajes dispersaron a Eurasia antes de 50 Ma. En consecuencia, la conquista del Holártico fue realizada por los antepasados de Hypericeae-Vismieae, antes de que tuviera lugar la primera diversificación de *Hypericum,* y presumiblemente no requirió ajustes fisiológicos para dispersarse a un Hemisferio Norte con condiciones tropicales.

13. Desde Eurasia linajes ancestrales de *Hypericum* entraron en Norte América antes de 35 Ma. Probablemente esta dispersión tuvo lugar a lo largo de Beringia, tal como revelan nuestras reconstrucciones basadas en modelos de paleodistribución de especies.
14. La cantidad de área con condiciones favorables para *Hypericum* en Beringia quedó drásticamente reducida después del Eoceno en comparación con otras áreas, probablemente interrumpiendo el flujo génico entre la poblaciones Holárticas de *Hypericum*.
15. La cantidad de área favorables para el desarrollo de *Hypericum* se redujo drásticamente en el este del Paleártico (EP) desde el Mioceno temprano hasta el presente, en comparación con otras regiones del Paleártico Occidentales y del Neártico, donde la cantidad de área favorable permaneció básicamente constante a lo largo del tiempo. Además, los linajes del EP presentan tasas de extinción y especiación más elevadas que los linajes de otroas regiones. Todo junto parece indicar que pudo haber existido extinción diferencial entre el EP y las otras regiones del planeta, lo que podría explicar la discrepancia encontrada entre el registro fósil y las reconstrucciones biogeográficas basadas en evidencias presentes.

16. Aunque las zonas templadas del Hemisferio Sur presentaron condiciones favorables para *Hypericum* a lo largo de toda su historia evolutiva, es muy probable que el cinturón tropical ecuatorial de estos continentes actuase como una barrera ecológica a la dispersión evitando que los linajes de *Hypericum* colonizaran las regiones más templadas del sur hasta el Mioceno Tardío. A partir de ese momento *Hypericum* probablemente dispersó al Hemisferio Sur a través de las recién formadas montañas de América del Sur, del este de África y Oeste de Asia.
17. Las tolerancias climáticas de *Hypericum* en el presente son comparables con aquellas que los ancestros poseían a finales del Eoceno, manteniéndose estables a lo largo de sus 35 Ma de historia evolutiva. Los resultados de este trabajo tienen implicaciones más allá de *Hypericum,* mostrando que las preferencias ecológicas de los organismos pueden preservarse durante periodos de tiempo de decenas de millones de años.
18. Aunque en su conjunto el nicho global del género se ha mantenido estable desde el Eoceno hasta el presente, *Hypericum* ha mantenido una tolerancia climática grande y

estable desde sus orígenes, el grupo también ha experimentado diversos eventos de evolución de nicho. Uno de los más significativos es aquel que ocurrió entre los ancestros de Hypericeae-Vismieae y el ancestro de todas lase species acutales *Hypericum* (crown group). Este cambio coincidió con el drástico descenso de las temperaturas que se produjo a finales del Eoceno y con la primera diversificación de *Hypericum*.

19. Los ancestros de Hypericaceae presentaban unas tolerancias intermedias entre clados tropicales y templados, adaptados a condiciones más frías pero más húmedas que clados tropicales. Las preferencias climáticas de estos ancestros posiblemente reflejan parte del nicho de un clado boreotropical.
20. Aunque en su conjunto el nicho global del género se ha mantenido estable desde el Eoceno hasta el presente, algunos clados se han especializado en distintas dimensiones del nicho.
21. En su conjunto, la evolución de las preferencias de *Hypericum* no ha sido direccional sino heterogénea en las diferentes dimensiones del nicho. Durante la evolución de *Hypericum* se puede observar una tendencia general hacia el aumento de la tolerancia al frío y a la falta de agua, por ejemplo en los clados *Hirtella* o *Hypericum*. Sin embargo, otros clados han mantenido unas condiciones más parecidas a la de los ancestros de *Hypericum* (*Elodes*, *Adenotrias* o *Myriandra*). Y algunos linajes han divergido significativamente de este patrón, como el grupo Afromontano que presenta tolerancias comparables a clados con afinidades tropicales, o *Ascyreia* que ha perdido la tolerancia a la falta de agua.
22. Los diferentes eventos de evolución de nicho en *Hypericum* no parece asociados a un aumento significativo de las tasas de diversificación. En su lugar, la edad del grupo puede por si sola explicar la riqueza de especies en *Hypericum*. *Hypericum* ha experimentado una acumulación constante de especies a lo largo del tiempo. Mientras que la evolución de nicho o de formas herbáceas ha podido contribuir a disminuir el riesgo de extinción en comparación con otros miembros del clado clusioide.
23. El éxito evolutivo de *Hypericum* radica en su elevada capacidad de resistencia antes el estress medioambiental, por medio de la dispersión y el cambio evolutivo. Esta resistencia posiblemente proviene de una gran plasticidad ancestral.

24. Esta tesis recalca la importancia que tiene un enfoque integral que al incorporar todas las fuentes de evidencia, tanto presentes como extintas, permite evitar el sesgo producido por la extinción en las reconstrucción filogenéticas. Este enfoque integral también permite escapar de las asunciones actualisticas que se utilizan de manera generalizada en biología evolutiva.

ANNEXES

1. ANNEXES CHAPTER 1

Appendix A

Voucher information and GenBank accession number for the taxa included in this study. The symbol * denotes sequences obtained from GenBank. The symbol ◆ denotes *psbA-trnH* sequences whose phylogenetic position was incongruent with those of other nuclear and plastid markers (see section 3.2).

Specie	Section	ID	Voucher	Locality	Genbank accession numbers			
					ITS	*trnL-trnF*	*psbA-trnH*	*trnS-trnG*
Hypericeae								
Hypericum aethiopicum subsp. aethiopicum Thunb	Adenosepalum	C116	GB1810 (GB)	South Africa, E. Cape Province	XX000000	XX000000	XX000000	XX000000
Hypericum aethiopicum subsp. sonderi (Bredell) N. Robson	Adenosepalum	C110	Aedo 14946 (MA)	South Africa, Orange Free State	XX000000	XX000000	XX000000	-
Hypericum aethiopicum subsp. sonderi (Bredell) N. Robson	Adenosepalum	C131	GB2057 (GB)	South Africa, Johanesburg	XX000000	XX000000	-	-
Hypericum afrum Lam.	Adenosepalum	C62	Dubuis s.n. (MA)	Algeria, Wilaya El Tarf	-	XX000000	-	-
Hypericum annulatum Moris	Adenosepalum	C7	Ryding 1485 (UPS)	Ethiopia, Eritrea	XX000000	XX000000	XX000000◆	-
Hypericum athoum Boiss. & Orp	Adenosepalum	AY555846	Crockett et al., 2004	-	AY555846*	-	-	-
Hypericum athoum Boiss. & Orp	Adenosepalum	C251	Sanchez 171 (MA)	Bot garden Goteborg	-	XX000000	XX000000	-
Hypericum atomarium Boiss	Adenosepalum	C267	Sorger 64-44-6 (W)	Turkey, Ephesus	-	-	XX000000	-
Hypericum caprifolium Boiss.	Adenosepalum	C17	Sanchez 3.1 (MA)	Spain, Tarragona	XX000000	XX000000	XX000000	-
Hypericum caprifolium Boiss.	Adenosepalum	C18	Sanchez 3.2 (MA)	Spain, Tarragona	XX000000	XX000000	XX000000	-
Hypericum coadunatum C. Smith ex Link	Adenosepalum	C144	Aldasoro A10353 (MA)	Spain, Gran Canaria	XX000000	XX000000	XX000000	XX000000
Hypericum conjungens N. Robson	Adenosepalum	C194	Mwasumbi 16191A (BM MO)	Tanzania, Mbeya	XX000000	XX000000	-	-
Hypericum conjungens N. Robson	Adenosepalum	C314	Mbago BG-Af 331 (Z)	Tanzania, Iringa	-	-	-	XX000000
Hypericum delphicum Boiss. & Heldr.	Adenosepalum	AY555845	Crockett et al., 2004	-	AY555845*	-	-	-
Hypericum delphicum Boiss. & Heldr.	Adenosepalum	FJ694197	Hazler-Pilepic & Blazina 2011	-	FJ694197*	-	-	-
Hypericum foliosum Aiton	Adenosepalum	C114	Aedo 10536 (MA)	Portugal, Azores, Isla Terceira	-	XX000000	XX000000	-
Hypericum glandulosum Aiton	Adenosepalum	C145	Aldasoro A10325 (MA)	Spain, Tenerife	XX000000	XX000000	XX000000	XX000000
Hypericum glandulosum Aiton	Adenosepalum	C150	Aldasoro A10349 (MA)	Spain, Tenerife	XX000000	XX000000	-	XX000000
Hypericum kiboënse Oliver	Adenosepalum	C1	Jonsell 2135 (UPS)	Tanzania, Kilimanjaro	-	XX000000	-	-
Hypericum kiboënse Oliver	Adenosepalum	C240	Sanchez 94 (MA)	Kenya, Kinangop, Aberdares Mts.	XX000000	XX000000	XX000000	XX000000
Hypericum kiboënse Oliver	Adenosepalum	C6	Hedberg 6350 (UPS)	Tanzania, Kitoto	XX000000	XX000000	-	XX000000
Hypericum lanuginosum Lam.	Adenosepalum	C127	Wok s.n. (GB)	Israel, Galilee	XX000000	-	-	-
Hypericum lanuginosum Lam.	Adenosepalum	C162	Haller s.n. (BC)	Israel, Nahal Qetalau	-	XX000000	-	-
Hypericum montanum L.	Adenosepalum	FJ694211	Hazler-Pilepic & Blazina 2011	-	FJ694211*	-	-	-
Hypericum montanum L.	Adenosepalum	C208	Aldasoro 14180 (MA)	Spain, Santander	XX000000	XX000000	XX000000	XX000000
Hypericum montanum L.	Adenosepalum	C90	Ferrero s.n. (MA)	Spain, Cuenca	XX000000	XX000000	-	XX000000
Hypericum naudinianum Coss. & Durieu	Adenosepalum	C157	Mateos 7107/95 (BC)	Morroco, Chefchaouen	-	XX000000	XX000000	-
Hypericum psilophytum (Diels) Maire	Adenosepalum	C38	Aldasoro A9867 (MA)	Algeria, Hoggar Mountains	XX000000	XX000000	XX000000	XX000000
Hypericum pubescens Boiss.	Adenosepalum	C104	Calvo JC1352 (MA)	Spain, Cadiz	XX000000	XX000000	XX000000	XX000000
Hypericum reflexum L. f.	Adenosepalum	C143	Aldasoro A10352 (MA)	Spain, Gran Canaria	XX000000	XX000000	XX000000	XX000000
Hypericum reflexum var. reflexum L. f.	Adenosepalum	C112	Marrero s.n. (MA)	Spain, Gran Canaria	XX000000	XX000000	-	-
Hypericum sinaicum Steudel & Hochst. ex Boiss.	Adenosepalum	C197	Danin 962609 (BM)	Jordan, Edom	XX000000	XX000000	XX000000◆	-

Hypericum somaliense N. Robson	Adenosepalum	C4	Thulin 9075 (UPS)	Somalia, Mirci	XX000000	XX000000	XX000000	-
Hypericum tomentosum L.	Adenosepalum	C19	Sanchez 4.1 (MA)	Spain, Tarragona	XX000000	XX000000	XX000000	-
Hypericum tomentosum L.	Adenosepalum	C20	Sanchez 4.2 (MA)	Spain, Tarragona	XX000000	XX000000	XX000000	-
Hypericum aegypticum L.	Adenotrias	C161	Di Martino s.n. (BC)	Italy, Sicilia	XX000000	XX000000	XX000000	XX000000
Hypericum aegypticum subsp. webbii (Spach) N. Robson	Adenotrias	C136	GB7706 (GB)	Greece, Santorini	XX000000	XX000000	XX000000	XX000000
Hypericum androsaemum L.	Androsaemum	FJ694190	Hazler-Pilepic & Blazina 2011	-	FJ694190*	-	-	-
Hypericum androsaemum L.	Androsaemum	C60	Sanchez 12 (MA)	Royal Bot garden Madrid	XX000000	XX000000	XX000000	XX000000
Hypericum grandifolium Choisy	Androsaemum	C146	Aldasoro A10354 (MA)	Spain, Gran Canaria	XX000000	XX000000	XX000000	XX000000
Hypericum grandifolium Choisy	Androsaemum	C147	Aldasoro A10316 (MA)	Spain, Tenerife	XX000000	XX000000	XX000000	XX000000
Hypericum hircinum L.	Androsaemum	FJ694204	Hazler-Pilepic & Blazina 2011	-	FJ694204*	-	-	-
Hypericum hircinum subsp. metroi L.	Androsaemum	C108	Calvo JC2576 (MA)	Morroco, Taza-Al	XX000000	XX000000	XX000000	XX000000
Hypericum x_inodorum Miller	Androsaemum	FJ694208	Hazler-Pilepic & Blazina 2011	-	FJ694208*	-	-	-
Hypericum pamphylicum N. Robson & P. Davis	Arthrophyllum	C196	Ulrich s.n. (BM)	Turkey, Antalya,	-	XX000000	XX000000	-
Hypericum acmosepalum N. Robson	Ascyreia	AY555851	Crockett et al., 2004	-	AY555851*	-	-	-
Hypericum acmosepalum N. Robson	Ascyreia	C302	Sino-British exp. Cangshan k052 (AAH)	China, W Yunnan	XX000000	XX000000	XX000000	-
Hypericum beanii N. Robson	Ascyreia	AY555852	Crockett et al., 2004	-	AY555852*	-	-	-
Hypericum beanii N. Robson	Ascyreia	C298	Sino-British exp. Cangshan K047 (AAH)	China, W Yunnan	-	XX000000	-	-
Hypericum calycinum L.	Ascyreia	AY555861	Crockett et al., 2004	-	AY555861*	-	-	-
Hypericum calycinum L.	Ascyreia	C58	Sanchez 10 (MA)	Royal Bot garden Madrid	XX000000	XX000000	XX000000	XX000000
Hypericum choisianum Wall. ex N. Robson	Ascyreia	AY555856	Crockett et al., 2004	-	AY555856*	-	-	-
Hypericum choisianum Wall. ex N. Robson	Ascyreia	C300	421 (AAH)	China, W Yunnan	-	XX000000	XX000000	-
Hypericum curvisepalum N. Robson	Ascyreia	C303	Bartholomew 120 (AAH)	China, W Yunnan	-	XX000000	XX000000	-
Hypericum dyeri Rehder	Ascyreia	C270	Steward 24528 (W)	Pakistan, Swat	XX000000	-	-	-
Hypericum elatoides Keller	Ascyreia	C305	Boufford 26156 (AAH)	China, Henan	-	XX000000	XX000000	XX000000
Hypericum forrestii (Chitt) Robson	Ascyreia	AY555858	Crockett et al., 2004	-	AY555858*	-	-	-
Hypericum forrestii (Chitt) Robson	Ascyreia	FJ694202	Hazler-Pilepic & Blazina 2011	-	FJ694202*	-	-	-
Hypericum forrestii (Chitt) Robson	Ascyreia	C296	Sino-British exp. Cangshan 423 (AAH)	China, W Yunnan	-	XX000000	-	-
Hypericum henryi H. Levl. & Van.	Ascyreia	C307	Li Heng 11347 (A)	China, Yunnan	XX000000	XX000000	XX000000	XX000000
Hypericum henryi subsp_uraloides (Rehder) N. Robson	Ascyreia	AY555859	Crockett et al., 2004	-	AY555859*	-	-	-
Hypericum hookerianum Wight & Arn.	Ascyreia	C284	Larsen 44980 (AAU)	Thailand, ChiangMai	-	XX000000	XX000000	XX000000
Hypericum hookerianum Wight & Arn.	Ascyreia	C309	Bartholomew 631 (A)	China, W Yunnan	XX000000	XX000000	-	-
Hypericum kouytchense H. Lév.	Ascyreia	AY555853	Crockett et al., 2004	-	AY555853*	-	-	-
Hypericum kouytchense H. Lév.	Ascyreia	FJ694210	Hazler-Pilepic & Blazina 2011	-	FJ694210*	-	-	-
Hypericum kouytchense H. Lév.	Ascyreia	FJ788906	Kosuth et al., 2010	-	-	-	FJ788906*	-
Hypericum lancasteri N. Robson	Ascyreia	AY555854	Crockett et al., 2004	-	AY555854*	-	-	-
Hypericum lancasteri N. Robson	Ascyreia	C299	Sino-British exp. Cangshan K047 (AAH)	China, W Yunnan	XX000000	XX000000	XX000000	-
Hypericum lancasteri N. Robson	Ascyreia	C311	Sino-British exp. Cangshan 1096 (A)	China, W Yunnan	-	XX000000	XX000000	XX000000
Hypericum leschenaultii Choisy	Ascyreia	AY555857	Crockett et al., 2004	-	AY555857*	-	-	-
Hypericum longistylum subsp. longistylum Oliver	Ascyreia	C301	Lancaster 1833 (AAH)	China, Hubei	XX000000	XX000000	XX000000	-
Hypericum monogynum L.	Ascyreia	C304	Lancaster 1828 (AAH)	China, E. Sichuan	XX000000	XX000000	-	-
Hypericum mysurense Wight & Arn.	Ascyreia	C286	Larsen 70-29605 (AAU)	Sri Lanka, Central Highlands	XX000000	-	-	-

Hypericum oblongifolium Choisy	Ascyreia	FJ694226	Hazler-Pilepic & Blazina 2011	-	FJ694226*	-	-	-
Hypericum oblongifolium Choisy	Ascyreia	C260	Ewald 6258 (GB)	Pakistan, Hazara	XX000000	-	-	-
Hypericum patulum Thunb. Ex Murray	Ascyreia	AY555860	Crockett et al., 2004	-	AY555860*	-	-	-
Hypericum patulum Thunb. Ex Murray	Ascyreia	C203	Aldasoro 14207 (MA)	Spain, Santander	XX000000	XX000000	XX000000	XX000000
Hypericum pseudohenryi N. Robson	Ascyreia	AY555850	Crockett et al., 2004	-	AY555850*	-	-	-
Hypericum pseudohenryi N. Robson	Ascyreia	C306	Boufford 32838 (AAH)	China, Sichuan	XX000000	XX000000	XX000000	XX000000
Hypericum subsessile N. Robson	Ascyreia	C308	Bartholomew 865 (A)	China, W Yunnan	XX000000	-	-	-
Hypericum wilsonii N. Robson	Ascyreia	FJ694225	Hazler-Pilepic & Blazina 2011	-	FJ694225*	-	-	-
Hypericum x_moserianum Luquet ex André	Ascyreia	AY555855	Crockett et al., 2004	-	AY555855*	-	-	-
Hypericum aciculare Kunth	Brathys	C262	Harling 13351 (GB)	Ecuador, Loja	XX000000	-	-	XX000000
Hypericum bryoides Gleason	Brathys	C68	Wood 4504 (MA)	Colombia, N Santander	XX000000	XX000000	-	-
Hypericum drummondii (Grev. & Hook) Torrey & Gray	Brathys	C119	Vicent 3958 (GB)	USA, Ohio	XX000000	XX000000	XX000000	-
Hypericum gentianoides (L) Britton	Brathys	C186	Miller 8429 (MO)	USA, Florida	XX000000	XX000000	XX000000	-
Hypericum juniperinum Kunth	Brathys	C83	Wood 4796 (MA)	Colombia, Cauca	XX000000	XX000000	-	-
Hypericum laricifolium Juss.	Brathys	C266	Persson 1622 (GB)	Ecuador, Pichinga	-	XX000000	XX000000	-
Hypericum laricifolium Juss.	Brathys	C316	Hilpold 10943 (BOZ)	Peru, Yungay	XX000000	XX000000	XX000000	-
Hypericum laricifolium Jussieu	Brathys	C263	Zak 3484 (GB)	Ecuador, Napo	XX000000	-	-	-
Hypericum mexicanum L.	Brathys	C86	Wood 5141 (MA)	Colombia, Boyaca	XX000000	XX000000	XX000000	XX000000
Hypericum pimelioides Planch. & Linden ex Triana &Planch.	Brathys	C102	Rangel 4025 (MA)	Colombia, Boyaca	-	XX000000	-	-
Hypericum quitense R. Keller	Brathys	C257	Antonelly 578 (GB)	Ecuador, Azuay	XX000000	XX000000	XX000000	XX000000
Hypericum sprucei N. Robson	Brathys	C265	Molau 3263 (GB)	Ecuador, Pichincha	XX000000	XX000000	-	-
Hypericum strictum Kunth	Brathys	C92	Brak s.n. (MA)	Costa Rica, Cartago	XX000000	XX000000	XX000000	-
Hypericum bupleuroides Griseb.	Bupleuroides	FJ788898	Kosuth et al., 2010	-	-	-	FJ788898*	-
Hypericum bupleuroides Griseb.	Bupleuroides	C65	Makaschrili s.n. (MA)	Georgia, Ajara	-	XX000000	-	-
Hypericum cerastoides (Spach) N. Robson	Campylopus	AY555884	Crockett et al., 2004	-	AY555884*	-	-	-
Hypericum cerastoides (Spach) N. Robson	Campylopus	C72	s.n. (MA)	Bulgaria, Kosovo	XX000000	XX000000	XX000000◆	XX000000
Hypericum balfourii N. Robson	Campylosporus	C171	Aldasoro 14697 (MA)	Yemen, Socotra	XX000000	XX000000	XX000000	XX000000
Hypericum bequaertii De Wild.	Campylosporus	C219	Sanchez 36 (MA)	Uganda, Rwenzori Mts.	XX000000	XX000000	XX000000	XX000000
Hypericum bequaertii De Wild.	Campylosporus	C220	Sanchez 38 (MA)	Uganda, Rwenzori Mts.	XX000000	XX000000	XX000000	XX000000
Hypericum dogonbadanicum Assadi	Campylosporus	C195	Assadi 38585 (BM)	Iran, Dogonbadan	XX000000	XX000000	XX000000	-
Hypericum quartinianum A. Rich	Campylosporus	C224	Sanchez 47 (MA)	Uganda, Kisumu, Mt. Elgon	XX000000	XX000000	XX000000	XX000000
Hypericum quartinianum A. Rich	Campylosporus	C32	Aldasoro A9986 (MA)	Ethiopia	XX000000	XX000000	XX000000	XX000000
Hypericum revolutum subsp. keniense (Scweinf.) N.Robson	Campylosporus	C215	Sanchez 32 (MA)	Uganda, Rwenzori Mts.	XX000000	XX000000	XX000000	XX000000
Hypericum revolutum subsp. revolutum Vahl (Schweinf)	Campylosporus	C213	Sanchez 28 (MA)	Uganda, Rwenzori Mts.	XX000000	XX000000	XX000000	XX000000
Hypericum revolutum Vahl (Schweinf)	Campylosporus	C82	Castroviejo 9145SC (MA)	Equatorial Guinea, Bioko	XX000000	XX000000	-	-
Hypericum roeperanum W. G. Schimper ex A. Rich	Campylosporus	C230	Sanchez 62 (MA)	Uganda, Kisumu, Mt. Elgon	XX000000	XX000000	XX000000	XX000000
Hypericum roeperanum W. G. Schimper ex A. Rich	Campylosporus	C233	Sanchez 70 (MA)	Uganda, Kisumu, Mt. Elgon	-	XX000000	XX000000	XX000000
Hypericum roeperanum W. G. Schimper ex A. Rich	Campylosporus	AY555863	Crockett et al., 2004	-	AY555863*	-	-	-
Hypericum socotranum subsp. socotranum Good	Campylosporus	C167	Aldasoro 14671 (MA)	Yemen, Socotra	XX000000	XX000000	XX000000	XX000000
Hypericum synstylum N. Robson	Campylosporus	C11	Burger 2422 (S)	Ethiopia, Harar prov.	XX000000	XX000000	XX000000◆	-
Hypericum synstylum N. Robson	Campylosporus	C3	Thulin 11038 (UPS)	Somalia,	XX000000	XX000000	XX000000	-

Hypericum amblycalyx Coust. & Gandoger	Coridium	C155	Curcó s.n. (BCN)	Greece, Creta	XX000000	-	-	-
Hypericum coris L.	Coridium	C23	Sanchez 5.1 (MA)	France, Alps Maritimes	XX000000	XX000000	XX000000◆	XX000000
Hypericum coris L.	Coridium	C24	Sanchez 5.2 (MA)	France, Alps Maritimes	XX000000	XX000000	XX000000	XX000000
Hypericum empetrifolium var. oliganthum Willd.	Coridium	C256	Sanchez 169 (GB)	Bot garden Goteborg	XX000000	-	XX000000	XX000000
Hypericum empetrifolium Willd.	Coridium	C200	Ruiz s.n. (MA)	Greece, atenas	XX000000	-	-	XX000000
Hypericum empetrifolium Willd.	Coridium	C255	Sanchez 168 (GB)	Bot garden Goteborg	XX000000	XX000000	XX000000	XX000000
Hypericum empetrifolium Willd.	Coridium	C70	Gadringer KRS5-6 (MA)	Greece, Creta	-	XX000000	-	-
Hypericum ericoides L.	Coridium	AY555847	Crockett et al., 2004	-	AY555847*	-	-	-
Hypericum ericoides L.	Coridium	C107	Calvo JC2308 (MA)	Spain, Albacete	XX000000	XX000000	XX000000	XX000000
Hypericum aucheri Jaub. & Spach	Crossophyllum	C37	Aldasoro A9794 (MA)	Turkey,	XX000000	XX000000	XX000000◆	XX000000
Hypericum orientale L.	Crossophyllum	FJ694213	Hazler-Pilepic & Blazina 2011	-	FJ694213*	-	-	-
Hypericum orientale L.	Crossophyllum	FJ788905	Kosuth et al., 2010	-	-	-	FJ788905*	-
Hypericum orientale L.	Crossophyllum	C244	Sanchez 166 (MA)	Bot garden Goteborg	-	XX000000	-	-
Hypericum barbatum Jacq.	Drosocarpium	FJ694192	Hazler-Pilepic & Blazina 2011	-	FJ694192*	-	-	-
Hypericum barbatum Jacq.	Drosocarpium	C118	s.n. (GB)	Bulgaria, Sofía	-	XX000000	-	-
Hypericum montbretii Spach	Drosocarpium	C84	Aedo 10350 (MA)	Bulgaria, Kosovo	XX000000	XX000000	XX000000	XX000000
Hypericum perfoliatum L.	Drosocarpium	C98	Aldasoro 3213 (MA)	Italy, Abruzzo	XX000000	XX000000	XX000000◆	XX000000
Hypericum richeri subsp. burseri (DC.) Nyman	Drosocarpium	C106	Romero s.n. (MA)	Spain, Leon	XX000000	XX000000	XX000000	XX000000
Hypericum richeri subsp. burseri (DC.) Nyman	Drosocarpium	C207	Aldasoro 14189 (MA)	Spain, Santander	XX000000	-	XX000000	-
Hypericum richeri subsp. grisebachii (Boiss.) Nyman	Drosocarpium	FJ694222	Hazler-Pilepic & Blazina 2011	-	FJ694222*	-	-	-
Hypericum rochelii Griseb. & Schenk	Drosocarpium	C95	Quintanar 1283AQ (MA)	Bulgaria, Blagoevgrad	XX000000	-	-	-
Hypericum rumeliacum Boiss.	Drosocarpium	C14	Emanuelsson 3001 (S)	Bulgaria, Asenovgrad	XX000000	XX000000	-	-
Hypericum elodeoides Choisy	Elodeoida	C135	Stainton 3562 (GB)	Nepal, Gurjakhani	XX000000	-	-	-
Hypericum elodes L.	Elodes	C166	Devain s.n. (MA)	Spain, Cantabria	XX000000	XX000000	XX000000	XX000000
Hypericum elodes L.	Elodes	C69	Peralta s.n. (MA)	Spain, Navarra	XX000000	XX000000	XX000000	XX000000
Hypericum graveolens Buckley	Graveolentia	AY555843	Crockett et al., 2004	-	AY555843*	-	-	-
Hypericum oaxacum Keller	Graveolentia	AY573003	Park & Kim 2004	-	AY573003*	-	-	-
Hypericum punctatum Lam.	Graveolentia	AY555844	Crockett et al., 2004	-	AY555844*	-	-	-
Hypericum punctatum Lam.	Graveolentia	GU562400	Fazekas et al., 2010	-	-	-	GU562400*	-
Hypericum heterophyllum Vent.	Heterophylla	C78	Nydegger 17659 (MA)	Turkey, Anatolia	XX000000	XX000000	-	-
Hypericum callithyrsum Coss.	Hirtella	C74	Pallares s.n. (MA)	Spain, Almeria	XX000000	XX000000	XX000000	XX000000
Hypericum helianthemoides (Spach) Boiss.	Hirtella	C77	Parisham s.n. (MA)	Iran, Isfahan	XX000000	XX000000	XX000000	XX000000
Hypericum hyssopifolium Vill.	Hirtella	C81	Medina LM2961 (MA)	Spain, Alava	XX000000	XX000000	XX000000	XX000000
Hypericum pseudolaeve N. Robson	Hirtella	C276	Sorger 82-71-10 (W)	Turkey, Karaagil	XX000000	XX000000	-	-
Hypericum scabrum L.	Hirtella	C91	Parisham s.n. (MA)	Iran, Isfahan	XX000000	XX000000	XX000000	-
Hypericum papuanum Ridl.	Humifusoideum	C275	Guilli 99 (W)	Papua New Guinea, E. Highlands	-	XX000000	XX000000	XX000000
Hypericum peplidifolium A. Rich	Humifusoideum	C26	Aldasoro A10057 (MA)	Ethiopia	XX000000	XX000000	XX000000	XX000000
Hypericum peplidifolium A. Rich	Humifusoideum	C27	Aldasoro A9971 (MA)	Ethiopia	XX000000	XX000000	XX000000	XX000000
Hypericum scioanum Chiov.	Humifusoideum	C29	Aldasoro A9957 (MA)	Ethiopia	XX000000	XX000000	XX000000	XX000000
Hypericum scioanum Chiov.	Humifusoideum	C30	Aldasoro A9991 (MA)	Ethiopia	XX000000	XX000000	XX000000	XX000000
Hypericum asahinae Makino	Hypericum	AY572997	Park & Kim 2004	-	AY572997*	-	-	-
Hypericum attenuatum Fisch. ex Choisy	Hypericum	AY572993	Park & Kim 2004	-	AY572993*	-	-	-
Hypericum attenuatum Fisch. ex Choisy	Hypericum	AY572995	Park & Kim 2004	-	AY572995*	-	-	-

Hypericum chejuense Park & Kim	Hypericum	AY572996	Park & Kim 2004	-	AY572996*	-	-	-
Hypericum elegans Stephan ex Willd.	Hypericum	C71	Cernoch s.n. (MA)	Bulgaria, Haskovo	XX000000	XX000000	XX000000	XX000000
Hypericum erectum Thunb. ex Murray	Hypericum	AY572991	Park & Kim 2004	-	AY572991*	-	-	-
Hypericum erectum Thunb. ex Murray	Hypericum	FJ788904	Kosuth et al., 2010	-	-	-	FJ788904*	-
Hypericum erectum Thunb. ex Murray	Hypericum	C202	García MAG 4071 (MA)	South Korea, Jeollabuk-do	XX000000	XX000000	XX000000	XX000000
Hypericum formosum Kunth.	Hypericum	C175	Merrill 12606 (MO)	USA, Colorado	XX000000	XX000000	XX000000	XX000000
Hypericum hakonense Franchet & Savat.	Hypericum	AY573000	Park & Kim 2004	-	AY573000*	-	-	-
Hypericum kamtschaticum Ledeb.	Hypericum	AY572992	Park & Kim 2004	-	AY572992*	-	-	-
Hypericum kamtschaticum Ledeb.	Hypericum	FJ793044	Hazler-Pilepic & Blazina 2011	-	FJ793044*	-	-	-
Hypericum kamtschaticum Ledeb.	Hypericum	83758492	Senni et al., 2005	-	-	83758492*	-	-
Hypericum kamtschaticum Ledeb.	Hypericum	83758494	Senni et al., 2005	-	-	83758494*	-	-
Hypericum kinashianum Koidz.	Hypericum	AY573001	Park & Kim 2004	-	AY573001*	-	-	-
Hypericum maculatum Crantz	Hypericum	C96	Aedo CA9479 (MA)	Andorra	XX000000	XX000000	XX000000	XX000000
Hypericum maculatum subsp. maculatum Crantz	Hypericum	C206	Aldasoro 14182 (MA)	Spain, Santander	XX000000	XX000000	XX000000	XX000000
Hypericum oliganthum Franchet & Savat.	Hypericum	AY573005	Park & Kim 2004	-	AY573005*	-	-	-
Hypericum ovalifolium Koidz.	Hypericum	AY572998	Park & Kim 2004	-	AY572998*	-	-	-
Hypericum perforatum L.	Hypericum	C178	Schmidt 1508 (MO)	USA, Pennsylvania	XX000000	XX000000	XX000000	XX000000
Hypericum perforatum L.	Hypericum	C22	Sanchez 1 (MA)	Spain, Tarragona	XX000000	XX000000	XX000000	XX000000
Hypericum perforatum L.	Hypericum	C47	Tauleigne s.n. (MA)	Portugal, Baixo Alentejo	XX000000	XX000000	XX000000	-
Hypericum perforatum L.	Hypericum	C56	Tauleigne s.n. (MA)	Portugal, Vinuoso	XX000000	XX000000	XX000000	-
Hypericum pseudopetiolatum Keller	Hypericum	AY573002	Park & Kim 2004	-	AY573002*	-	-	-
Hypericum scouleri Hook.	Hypericum	C80	Twisselmann 11364 (MA)	USA, Tulare	XX000000	XX000000	XX000000	-
Hypericum sikokumontanum Makino	Hypericum	AY572999	Park & Kim 2004	-	AY572999*	-	-	-
Hypericum tetrapterum Fries	Hypericum	FJ694224	Hazler-Pilepic & Blazina 2011		FJ694224*	-	-	-
Hypericum tetrapterum Fries	Hypericum	FJ788897	Kosuth et al., 2010	-	-	-	FJ788897*	-
Hypericum tetrapterum Fries	Hypericum	C21	Sanchez 2 (MA)	Spain, Tarragona	XX000000	XX000000	XX000000◆	XX000000
Hypericum triquetrifolium Turra	Hypericum	C39	Aldasoro A9795 (MA)	Turkey	XX000000	XX000000	XX000000◆	XX000000
Hypericum undulatum Schousboe ex Willd.	Hypericum	C156	Vigo s.n. (BCN)	Spain, Soria	-	XX000000	-	-
Hypericum undulatum Schousboe ex Willd.	Hypericum	C99	Serra 6034 (MA)	Spain, Oviedo	XX000000	XX000000	XX000000◆	XX000000
Hypericum vaniotii Lev.	Hypericum	AY572994	Park & Kim 2004	-	AY572994*	-	-	-
Hypericum yezoense Maxim.	Hypericum	FJ793046	Hazler-Pilepic & Blazina 2011	-	FJ793046*	-	-	-
Hypericum yezoense Maxim.	Hypericum	AY573004	Park & Kim 2004	-	AY573004*	-	-	-
Hypericum xylosteifolium (Spach) N. Robson	Inodora	C274	Sorger 69-23-28 (W)	Turkey, Steilhange	XX000000	-	XX000000	XX000000
Hypericum monanthemum Hook. F. & Thomsom ex Dyer	Monanthema	C283	Larsen 46519 (AAU)	Thailand, ChiangMai	XX000000	XX000000	-	-
Hypericum wightianum Wall.	Monanthema	C273	Kingdom-Ward 22448 (W)	Burma, Mindat	-	-	XX000000	XX000000
Hypericum adpressum W. Barton	Myriandra	AY555865	Crockett et al., 2004	-	AY555865*	-	-	-
Hypericum adpressum W. Barton	Myriandra	C137	P.O.S. 21696 (GB)	USA, Florida	XX000000	XX000000	XX000000	XX000000
Hypericum apocynifolium Small	Myriandra	AY555883	Crockett et al., 2004	-	AY555883*	-	-	-
Hypericum brachyphyllum (Spach) Steud.	Myriandra	AY555870	Crockett et al., 2004	-	AY555870*	-	-	-
Hypericum brachyphyllum (Spach) Steud.	Myriandra	C181	Miller 8438 (MO)	USA, Florida	XX000000	XX000000	XX000000	XX000000
Hypericum buckleyi Curtis	Myriandra	AY555880	Crockett et al., 2004	-	AY555880*	-	-	-
Hypericum chapmanii Adams	Myriandra	AY555869	Crockett et al., 2004	-	AY555869*	-	-	-
Hypericum cistifolium Lam.	Myriandra	C176	Bradley 1186 (MO)	USA, Florida	XX000000	XX000000	XX000000	XX000000

Hypericum cistifolium Lam.	Myriandra	C177	Miller 8393 (MO)	USA, Florida	XX000000	XX000000	XX000000	-
Hypericum crux-andreae (L) Crantz	Myriandra	AY555874	Crockett et al., 2004	-	AY555874*	-	-	-
Hypericum crux-andreae (L) Crantz	Myriandra	C174	Miller 8455 (MO)	USA, Florida	XX000000	XX000000	XX000000	-
Hypericum densiflorum Pursh	Myriandra	AY555886	Crockett et al., 2004	-	AY555886*	-	-	-
Hypericum densiflorum Pursh	Myriandra	C73	Thomas 97505 (MA)	USA, Ashley	-	XX000000	-	-
Hypericum dolabriforme Vent.	Myriandra	AY555889	Crockett et al., 2004	-	AY555889*	-	-	-
Hypericum ellipticum Hook.	Myriandra	C5	Schepanek 6623 (UPS)	Canada, McAdam Parish	XX000000	XX000000	XX000000	XX000000
Hypericum fasciculatum Lam.	Myriandra	AY555868	Crockett et al., 2004	-	AY555868*	-	-	-
Hypericum fasciculatum Lam.	Myriandra	C173	Bradley 1187 (MO)	USA, Florida	XX000000	XX000000	XX000000	XX000000
Hypericum fasciculatum Lam.	Myriandra	C76	Carrasco s.n. (MA)	Cuba, Santiago de Cuba	XX000000	XX000000	XX000000	XX000000
Hypericum frondosum Michaux	Myriandra	AY555887	Crockett et al., 2004	-	AY555887*	-	-	-
Hypericum galioides Lam.	Myriandra	AY555864	Crockett et al., 2004	-	AY555864*	-	-	-
Hypericum galioides Lam.	Myriandra	C133	Boufford 5149 (GB)	Georgia, Evans	XX000000	XX000000	XX000000	XX000000
Hypericum hypericoides (L.) Crantz	Myriandra	C132	Vicent 4291 (GB)	USA, N. Carolina, Union	XX000000	XX000000	XX000000	XX000000
Hypericum hypericoides (L.) Crantz	Myriandra	C185	Miller 8447 (MO)	USA, Florida	XX000000	XX000000	XX000000	XX000000
Hypericum kalmianum L.	Myriandra	FJ694209	Hazler-Pilepic & Blazina 2011	-	FJ694209*	-	-	-
Hypericum kalmianum L.	Myriandra	FJ788896	Kosuth et al., 2010	-	-	-	FJ788896*	-
Hypericum lissophloeus P. Adams	Myriandra	AY555885	Crockett et al., 2004	-	AY555885*	-	-	-
Hypericum lissophloeus P. Adams	Myriandra	C134	Godfrey 61554 (GB)	USA, Florida, Bay	XX000000	XX000000	-	XX000000
Hypericum lloydii (Svenson) P. Adams	Myriandra	AY555867	Crockett et al., 2004	-	AY555867*	-	-	-
Hypericum lobocarpum Gattinger	Myriandra	AY555876	Crockett et al., 2004	-	AY555876*	-	-	-
Hypericum microsepalum (Torrey & Gray) Gray ex Watson	Myriandra	AY555877	Crockett et al., 2004	-	AY555877*	-	-	-
Hypericum myrtifolium Lam.	Myriandra	AY555875	Crockett et al., 2004	-	AY555875*	-	-	-
Hypericum nitidum Lam.	Myriandra	AY555871	Crockett et al., 2004	-	AY555871*	-	-	-
Hypericum nudiflorum Michaux	Myriandra	AY555888	Crockett et al., 2004	-	AY555888*	-	-	-
Hypericum prolificum L.	Myriandra	C182	Nye 243 (MO)	USA, Missouri	XX000000	XX000000	XX000000	XX000000
Hypericum prolificum L.	Myriandra	C97	Ahles 87220 (MA)	USA, Massachuset	XX000000	XX000000	XX000000	XX000000
Hypericum sphaerocarpum Michaux	Myriandra	AY555878	Crockett et al., 2004	-	AY555878*	-	-	-
Hypericum tenuifolium Pursh	Myriandra	AY555872	Crockett et al., 2004	-	AY555872*	-	-	-
Hypericum tenuifolium Pursh	Myriandra	C13	Bradley 3345 (S)	USA, North Carolina	XX000000	XX000000	XX000000	XX000000
Hypericum tetrapetalum Lam.	Myriandra	AY555882	Crockett et al., 2004	-	AY555882*	-	-	-
Hypericum tetrapetalum Lam.	Myriandra	C122	Vicent 5153 (GB)	USA, Florida, Levy	XX000000	XX000000	XX000000	-
Hypericum humifusum L.	Oligostema	FJ788903	Kosuth et al., 2010	-	-	-	FJ788903*	-
Hypericum humifusum L.	Oligostema	C201	Ruiz s.n. (MA)	Morroco, Tetuan	XX000000	XX000000	XX000000	-
Hypericum humifusum L.	Oligostema	C204	Aldasoro 14208 (MA)	Spain, Santander	XX000000	XX000000	XX000000	-
Hypericum linariifolium Vahl	Oligostema	C63	Amaraz s.n. (MA)	Spain, Cáceres	XX000000	XX000000	-	-
Hypericum linariifolium Vahl	Oligostema	C165	Gómiz s.n. (BC)	Spain, Leon	XX000000	XX000000	XX000000	-
Hypericum linariifolium Vahl	Oligostema	C46	Tauleigne s.n. (MA)	Portugal, Baixo Alentejo	XX000000	XX000000	XX000000	XX000000
Hypericum olympicum L.	Olympia	C199	Ruiz s.n. (MA)	Greece, Laconia	XX000000	XX000000	XX000000	XX000000
Hypericum olympicum L.	Olympia	C57	Sanchez AS9 (MA)	Royal Bot garden Madrid	XX000000	XX000000	XX000000	XX000000
Hypericum polyphyllum Boiss. & Balansa	Olympia	FJ694216	Hazler-Pilepic & Blazina 2011	-	FJ694216*	-	-	-
Hypericum balearicum L.	Psorophytum	C40	Saez 5006 (MA)	Spain, Mallorca	XX000000	XX000000	XX000000◆	XX000000
Hypericum balearicum L.	Psorophytum	C61	Sanchez 13 (MA)	Royal Bot garden Madrid	XX000000	XX000000	XX000000	XX000000
Hypericum ascyron L.	Roscyna	FJ694189	Hazler-Pilepic & Blazina 2011	-	FJ694189*	-	-	-

Hypericum ascyron subsp ascyron L.	Roscyna	C44	MAGarcía 4059 (MA)	South Korea, Jeollakbuk-do	XX000000	XX000000	XX000000	XX000000
Hypericum sampsonii Hance	Sampsonia	AY573011	Park & Kim 2004	-	AY573011*	-	-	-
Hypericum confertum Choisy	Taeniocarpium	C138	Lindberg s.n. (GB)	Cyprus, Mt. Troodos	XX000000	XX000000	XX000000	XX000000
Hypericum hirsutum L.	Taeniocarpium	FJ694203	Hazler-Pilepic & Blazina 2011	-	FJ694203*	-	-	-
Hypericum hirsutum L.	Taeniocarpium	C59	Sanchez 11 (MA)	Royal Bot garden Madrid	XX000000	XX000000	XX000000	XX000000
Hypericum linarioides Bosse	Taeniocarpium	C88	Aldasoro 2667 (MA)	Turkey, Sakaltutan	XX000000	XX000000	XX000000	XX000000
Hypericum nummularioides Trautv.	Taeniocarpium	C243	Sanchez 164 (MA)	Bot garden Goteborg	-	XX000000	XX000000	-
Hypericum nummularium L.	Taeniocarpium	C101	Jauregui s.n. (MA)	Spain, Navarra	XX000000	XX000000	XX000000	-
Hypericum nummularium L.	Taeniocarpium	C205	Aldasoro 14179 (MA)	Spain, Santander	XX000000	XX000000	XX000000	XX000000
Hypericum pulchrum L.	Taeniocarpium	FJ694219	Hazler-Pilepic & Blazina 2011	-	FJ694219*	-	-	-
Hypericum venustum Fengl	Taeniocarpium	C280	Sorger 81-27-21 (W)	Turkey, Hakkari	-	XX000000	-	-
Hypericum geminiflorum Hemsley	Takasagoya	C12	Chung 1266 (S)	China, Taiwan, Pingtung Hsien	HM162838	-	-	-
Hypericum thasium Griseb.	Thasia	C278	Rechinger 45280 (W)	Greece, thasos	-	XX000000	-	-
Hypericum pallens Banks & Solander	Triadenioides	C253	Sanchez 167 (MA)	Bot garden Goteborg	-	-	XX000000	-
Hypericum pallens Banks & Solander	Triadenioides	AY555848	Crockett et al., 2004	-	AY555848*	-	-	-
Hypericum scopulorum Balf. f.	Triadenioides	C169	Aldasoro 14644 (MA)	Yemen, Socotra, Magarhar	XX000000	XX000000	XX000000	XX000000
Hypericum tortuosum Balf. f.	Triadenioides	C170	Aldasoro 14645 (MA)	Yemen, Socotra	XX000000	XX000000	XX000000	XX000000
Hypericum boreale (Britton) Bickn.	Trigynobrathys	AY573026	Park & Kim 2004	-	AY573026*	-	-	-
Hypericum boreale (Britton) Bickn.	Trigynobrathys	C130	Ahles 86328 (GB)	USA, Massachuset	XX000000	XX000000	XX000000	XX000000
Hypericum brevistylum Choisy	Trigynobrathys	AY573019	Park & Kim 2004	-	AY573019*	-	-	-
Hypericum brevistylum Choisy.	Trigynobrathys	C317	Hilpold 11745 (BOZ)	Peru, Cuzco	XX000000	XX000000	-	-
Hypericum brevistylum Choisy	Trigynobrathys	C318	Hilpold 11413 (BOZ)	Peru, Ancash	XX000000	XX000000	-	-
Hypericum canadense L.	Trigynobrathys	C259	Brisson 12774 (GB)	Canada, Lac Aylmer	XX000000	XX000000	-	XX000000
Hypericum gramineum G. Foster	Trigynobrathys	EU352256	Heenan 2008	-	EU352256*	-	-	-
Hypericum gramineum G. Foster	Trigynobrathys	EU352257	Heenan 2008	-	EU352257*	-	-	-
Hypericum japonicum Thunb. ex Murray	Trigynobrathys	AY573025	Park & Kim 2004	-	AY573025*	-	-	-
Hypericum japonicum Thunb. ex Murray	Trigynobrathys	FJ980417	Chen & Han, unpublish	-	FJ980417*	-	-	-
Hypericum japonicum Thunb. ex Murray	Trigynobrathys	GQ435379	Chen et al., 2010	-	-	-	GQ435379*	-
Hypericum jeongjocksanense Park & Kim	Trigynobrathys	AY573023	Park & Kim 2004	-	AY573023*	-	-	-
Hypericum lalandii Choisy	Trigynobrathys	C128	Dahlstrand 2633 (GB)	South Africa, E. Cape Provice	XX000000	XX000000	XX000000	-
Hypericum lalandii Choisy	Trigynobrathys	C248	Dahlstrand 1102 (GB)	South Africa, Transvaal	XX000000	XX000000	-	-
Hypericum laxum (Bl.) Koidz.	Trigynobrathys	AY573024	Park & Kim 2004	-	AY573024*	-	-	-
Hypericum majus (A. Gray) Britton	Trigynobrathys	C85	Rastetter s.n. (MA)	France, Haute-Saone	XX000000	XX000000	XX000000	-
Hypericum mutilum L.	Trigynobrathys	DQ006195	Kress et al., 2005	-	-	-	DQ006195*	-
Hypericum mutilum L.	Trigynobrathys	C164	Lazare s.n. (BC)	France, Landes	XX000000	XX000000	XX000000	XX000000
Hypericum mutilum subsp. boreale (Britton) J. M. Gillett	Trigynobrathys	C179	Schmidt 1488 (MO)	USA, Ohio	XX000000	XX000000	XX000000	XX000000
Hypericum myrianthum subsp. tamariscinum (C&S) Robson	Trigynobrathys	C264	Pedersen 15904 (GB)	Brasil, Restinga Seca	XX000000	-	-	-
Hypericum rigidum A. St. Hil.	Trigynobrathys	AY573021	Park & Kim 2004	-	AY573021*	-	-	-
Hypericum setosum L.	Trigynobrathys	AY573020	Park & Kim 2004	-	AY573020*	-	-	-
Hypericum silenoides subsp. silenoides Juss.	Trigynobrathys	C67	Basualto (MA)	Chile, VIII region, Concepcion	XX000000	XX000000	XX000000	-
Hypericum ternum A. St. Hil.	Trigynobrathys	AY573022	Park & Kim 2004	-	AY573022*	-	-	-
Hypericum canariense L.	Webbia	C148	Aldasoro A10304 (MA)	Spain, Tenerife	XX000000	XX000000	XX000000	XX000000
Hypericum canariense L.	Webbia	C151	Aldasoro A10312 (MA)	Spain, Tenerife	XX000000	XX000000	XX000000	XX000000

Thornea calcicola (Standl. & Steyerm.) Breedl. & McClintock	-	AY573028	Park & Kim 2004	-	AY573028*	-	-	-
Thornea matudae (Lundell) Breedl. & McClintock	-	AY573027	Park & Kim 2004	-	AY573027*	-	-	-
Triadenum fraseri (Spach) Gleason	-	C282	Ford 547 (W)	Canada, Manitoba	XX000000	XX000000	XX000000	XX000000
Triadenum petiolatum Hook f. & Thomson ex Dyer	-	C16	Correll 35026 (S)	USA, Texas	XX000000	XX000000	XX000000◆	-
Vismieae								
Harungana madagascarensis Lam. ex Poir.	-	C105	Fernandez Casas s.n. (MA)	Equatorial Guinea, Bioko	XX000000	XX000000	XX000000	XX000000
Psorospermum senegalense Spach	-	C109	Duvale 549 (MA)	Mali, Korofing National Park	-	XX000000	XX000000	-
Vismia glaziovii Ruhland	-	C190	Fuentes 10934 (MO)	Bolivia, La Paz	XX000000	XX000000	XX000000	XX000000
Vismia rubescens Oliv.	-	C192	Niangadouma 374 (MO)	Gabon, Haute-Ogooue	XX000000	-	XX000000	-
Cratoxyleae								
Eliea articulata (Lam.) Cambess	-	C189	Razakamalala 295 (MO)	Madagascar, Fianarantsoa	XX000000	XX000000	XX000000	-

SI Appendix

I. Characters coded for the Bayesian ancestral state reconstruction.

1/ Habit form with three states (tree, shrub, herb),

2/ presence/absence of dark glands in vegetative and reproductive tissues,

3/ number of vestigial fascicles in the androecia (0, 3, 5),

4/ sculpturing pattern of the seed testa (scalariform, reticulate, papillose, smooth),

5/ shape of the corolla (stellate, cyathiform and pseudo-tubular),

6/ number of stamen fascicles (3, 5),

7/ fusion of stamen fascicles (fused, partly fussed, free). Robson (1981) defined 10 androecial configurations in *Hypericum* that varied from five free fascicles to totally united fascicles forming a narrow or broad continuous ring. Intermediate states appear as a consequence of the condensation of fascicles and or elimination of some of them. As defined by Robson, the limits between states in this character are sometimes not clearly recognized - for example between the union of fascicles to form a narrow continuous ring (character state 'e' in Robson, 1981) or the merging of fascicles with partial obscuring of members (character state 'f' in Robson, 1981). Therefore, to reduce the ambiguity in the classification of specimens we created only 3 categories for this trait: free, fused and partially fused. The difference between the last two categories is based in whether the fascicles are in contact with each other to form a continuous ring or not.

8/ present distribution range (Afrotropical "AF", Eastern Palearctic "EP", ‚Irano–Turanian-Himalayan region "IT"Nearctic "NE", Neotropical "NT", Oceania "OC", and Western Palearctic "WP").

II. Delimitation of areas in the biogeographic analyses.

We based our delimitation of areas on both patterns of endemicity in *Hypericum* (shared distributions among species) and the paleogeographic history of the continents. Some of our areas are congruent with the floristic regions of Takhtajan (Takhtajan, 1986), which was based on patterns of vegetation endemism, and therefore likely reflect geographical and eco-physiographical (climatical) features, cf. Takhtajan, (Takhtajan, 1986). We also tried to maximize congruence with other biogeographic studies (Sanmartín et al., 2001).

Seven operational areas were defined (Fig. 5):

1) Western Palaearctic ("WP"): Including Europe, Macaronesia, and Northern Africa (north of the Saharan belt). The latter was formed by the collision of Africa with the Eurasian plate, and the accretion of several terranes (Meulenkamp and Sissingh, 2003; Rosenbaum et al., 2002). Takhtajan (Takhtajan, 1986) considered North Africa as part of the Mediterranean floristic region. The WP region is limited in the east by the former Turgai Strait that divided the Palearctic until the Early Oligocene (Sanmartín et al., 2001).
2) Eastern Palaearctic region ("EP"): Includes non-tropical Asia east of the Turgai Strait as defined by Sanmartín et al. (Sanmartín et al., 2001), but excluding area "ITH" (see below).
3) Irano-Turanian-Himalayan region ("ITH"): includes all the new-uplifted landmasses and mountain ranges that resulted from the collision of the Indian and Arabian plates against Eurasia in the Eocene and Middle Miocene, respectively (Sanmartín, 2003; Wang et al., 2009). Before the Tertiary, this area formed the eastern arm of the Tethys Seaway. The ITS region extends from the Zagros Mountains and the Caucasus in the west through the Iranian plateau, the Hindu Kush, Tien Shan and Kunlun Shan Mountains to the Altay Mountains in the northeast, and the Tibetan Plateau and the Himalayan Mountain range in the southeast.
4) Nearctic region ("NE"): Includes North America, Mexico and the Caribbean islands. We included Mexico here down to the Mexican lowlands because it was connected to Western North America via the Cordillera Mountain System, including the Rocky Mountains and Sierra Madre Occidental (Sanmartín et al., 2001). The Caribbean region was included in the Nearctic because there are very few endemics in this region and most widespread species are shared with eastern North America.
5) Oceania ("OC"): Includes the tropical parts of South East Asia that were formed by the accretion of Gondwanan terranes to Asia at different times during the Late Mesozoic – Cenozoic (Metcalfe, 1996; Sanmartín and Ronquist, 2004) – the tropical regions of southwest and southern China, eastern parts of Bangladesh, Burma, Thailand, Indochina, the island of Hainan and the Malaysian Peninsula – as well as the Indian subcontinent (all Indostan excluding the tropical foothills of Himalayas in the north), Australia, New Zealand, New Caledonia, the Polynesian region, and Hawaii.
6) Africa ("AF"): Including the region south of the Saharan belt, the island of Socotra, the Arabian Peninsula, Madagascar and the Mascarene Islands.
7) Neotropics ("NT"): Including the South American continent and Central America south from the Mexican lowlands.

There are seven terminals that represent widespread taxa distributed in more than one region. They were coded for the voucher original area for the analysis shown in Figure 5, but with the alternative area in a second analysis as follows: *H. ascyron*_C44 (NE); *Vismia*_C190 (NT); *Triadenum*_C282 (EP); *H. annulatum*_C7 (WP); *H. mutilum*_C179 (OC); *H. mutilum*_C164 (EP); *H. perforatum*_C56 (EP).

III. Phylogenetic position of the fossil *H. antiquum*

Meseguer & Sanmartín (Meseguer and Sanmartín, 2012) ascribed this fossil remain to the genus *Hypericum* based on the seed size, colour, shape and sculpturing pattern of the seed testa. The later character is particularly distinctive, as it is highly variable between *Hypericum* species, and sometimes allows distinguishing among sections (Meseguer and Sanmartín, 2012; Robson, 1981). *Hypericum antiquum* (Balueva & Nikitin) was found in an Upper Eocene site in West Siberia, and exhibits a ribbed-scalariform seed surface. This sculpturing pattern is present in several *Hypericum* sections: *Elodes, Brathys, Trigynobrathys,* and *Drosocarpium*, but other sections present the most common reticulate seed testa.

Results from our Bayesian ancestral state reconstruction for the pattern of the seed testa (Fig. 4) show that the ancestor of living *Hypericum* species probably possessed a "reticulate" pattern in the seed surface, which is also the most common state shared with the rest of tribes (Fig. 4). The scalariform seed testa was not inferred as the ancestral state of any of the main lineages, but appears scattered in several clades in the phylogeny (e.g., clades A, B, E). If the presence of the scalarifom seed testa in unrelated sections of *Hypericum* is convergent, and has evolved independently as suggested by our results (Fig. 4), then we could have assigned the fossil to the crown node of one of those clades. This would have pushed down the age of *Hypericum*, making it older. However, we do not have any morphological evidence to assign this fossil to a particular section. Magallón & Sanderson (Magallón and Sanderson, 2001) suggested that when a fossil presents a synapomorphy characteristic of a particular clade within the study group, it should be assigned to the crown node of the group; if the fossil presents some but not all the synapomorphies of the group, it should be assigned to the stem node. In consequence, we assigned the fossil to the crown node of *Hypericum*, in what we believe is a more conservative approach (the MRCA of all clades with scalariform seed testa corresponds to the crown-node in our phylogeny). Finally, there is the possibility that

differential extinction, as discussed in the text (see discussion), might have obscured the phylogenetic signal, preventing us from recovering the true ancestral character state for the crown node.

IV. Detailed results from the divergence time analysis

Node age of Hypericaceae and major subclades in million years. Mean, lower (Minimum) and upper (Maximum) values of the 95 % highest posterior density intervals of the posterior probability of distribution of node ages from the divergence times analysis run in BEAST.

Clade	Min	Mean	Max
Hypericaceae	43,77	53,87	66,47
Vismieae+Hypericeae	41,19	49,90	60,88
Vismieae	14,19	24,66	36,91
Hypericeae(=Hypericum)	34,01	34,92	37,22
Clade A	14,89	27,72	35,23
Clade B+C+D+E	30,30	33,75	36,78
Clade B	23,82	29,3	33,67
Brathys+Myriandra	16,99	21,92	27,33
Myriandra	9,41	13,59	18,48
Brathys group	9,19	13,68	18,47
Clade C+D+E	22,22	27,16	31,80
Clade C	12,62	21,02	28,34
Clade D+E	19,61	24,39	29,19
Clade D	10,57	15,99	21,75
Afromontane group	1,66	4,01	7,92
Ascyreia group	8,52	12,33	17,04
Clade E	14,76	18,85	23,19
Hirtella group	10,72	14,84	19,32
Adenosepalum group	8,24	11,39	14,91
Hypericum group	7,99	10,83	14,06
Adenosp.+Hyper. groups	10,04	13,31	16,88

SI Bibliography

Magallón, S., Sanderson, M.J., 2001. Absolute diversification rates in angiosperm clades. Evolution 55, 1762-1780.

Meseguer, A.S., Sanmartín, I., 2012. Paleobiology of the genus Hypericum (Hypericaceae): A survey of the fossil record and its palaeogeographic implications. Anales del Jardin Botanico de Madrid 69, 97-106.

Metcalfe, I., 1996. Pre-Cretaceous evolution of SE Asian terranes. In: Hall, R., Blundell, D.J. (Eds.), Tectonic evolution of South-east Asia. Geological Society of America, Boulder, Colorado, pp. 97–122.

Meulenkamp, J.E., Sissingh, W., 2003. Tertiary palaeogeography and tectonostratigraphic evolution of the Northern and Southern Peri-Tethys platforms and the intermediate domains of the African-Eurasian convergent plate boundary zone. Palaeogeography, Palaeoclimatology, Palaeoecology 196, 209-228.

Robson, N.K.B., 1981. Studies in the genus *Hypericum* L. (Guttiferae). 2. Characters of the genus. Bull. Brit. Mus. (Nat. Hist.), Bot. 8, 55–226.
Rosenbaum, G., Lister, G.S., Duboz, C., 2002. Reconstruction of the tectonic evolution of the western Mediterranean since the Oligocene. Journal of the Virtual Explorer 8, 107-130.
Sanmartín, I., 2003. Dispersal vs. vicariance in the Mediterranean: Historical biogeography of the Palearctic Pachydeminae (Coleoptera, Scarabaeoidea). Journal of Biogeography 30, 1883-1897.
Sanmartín, I., Enghoff, H., Ronquist, F., 2001. Patterns of animal dispersal, vicariance and diversification in the Holarctic. Biological Journal of the Linnean Society 73, 345-390.
Sanmartín, I., Ronquist, F., 2004. Southern Hemisphere biogeography inferred by event-based models: Plant versus animal patterns. Systematic Biology 53, 216-243.
Takhtajan, A., 1986. Floristic regions of the world. University of California Press, Berkeley.
Wang, Y.J., Susanna, A., Von Raab-Straube, E., Milne, R., Liu, J.Q., 2009. Island-like radiation of Saussurea (Asteraceae: Cardueae) triggered by uplifts of the Qinghai-Tibetan Plateau. Biological Journal of the Linnean Society 97, 893-903.

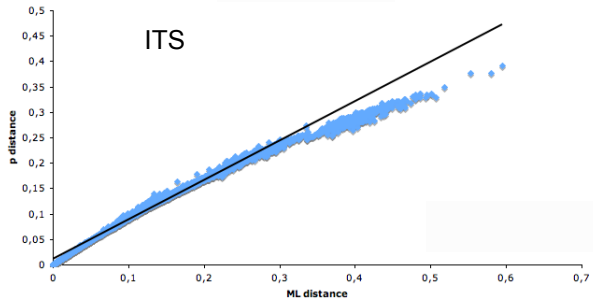
Supplementary information (SI) figure captions

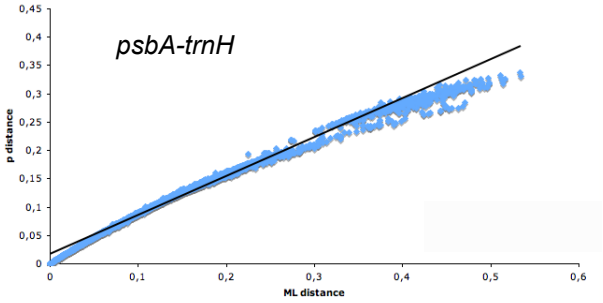
SI Figure 1. Saturation plots of the single-gene nuclear and chloroplast markers. The plot shows uncorrected ("p") pairwise distances against corrected maximum-likelihood distances derived in PAUP with the selected model of substitution.

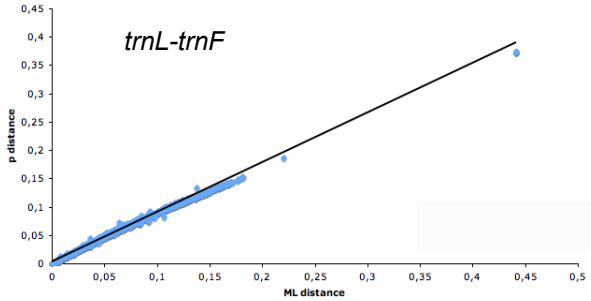
SI Figure 2. **Phylogenetic relationships in *Hypericum* inferred from the single-gene plastid markers.** The figure shows the 50% majority-rule consensus tree from an MCMC Bayesian analysis of: a) *psbA-trnH*, b) *trnL-trnF*, c) *trnS-trnG*. Bayesian posterior probabilities (pp) are indicated below branches; bootstrap support values for maximum likelihood (ML) rearrangements (500 replicates in GARLI) are shown above branches. Specimens found in an incongruent position between *psbA-trnH* and the other study markers are marked in blue.

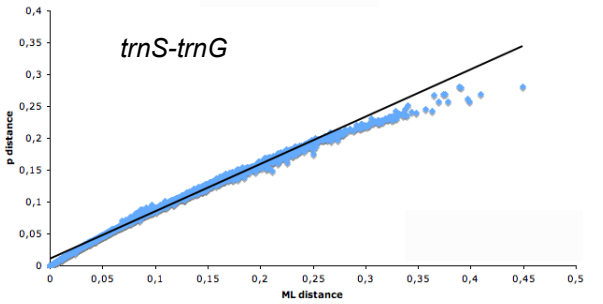
SI Figure 3. **Phylogenetic relationships in *Hypericum* inferred from the concatenated "All-specimens" plastid dataset including incongruent specimens on *psbA-trnH*.** Bayesian Majority-Rule consensus tree of *Hypericum* inferred from concatenated chloroplast markers (*psbA-trnH*, *trnL-trnF*, *trnS-trnG*), including incongruent specimens on *psbA-trnH*. Bootstrap (bp) support values are indicated above branches for ML rearrangements (500 replicates in GARLI) and Bayesian posterior probabilities below branches (pp). The inclusion of the incongruent specimens on *psbA-trnH* mainly affected the topology of clade E, e.g., *H. cerastoides* and *H. orientale* appear as sister group of the remaining species of the clade, instead of as sister to section *Oligostema* like in Fig. 3

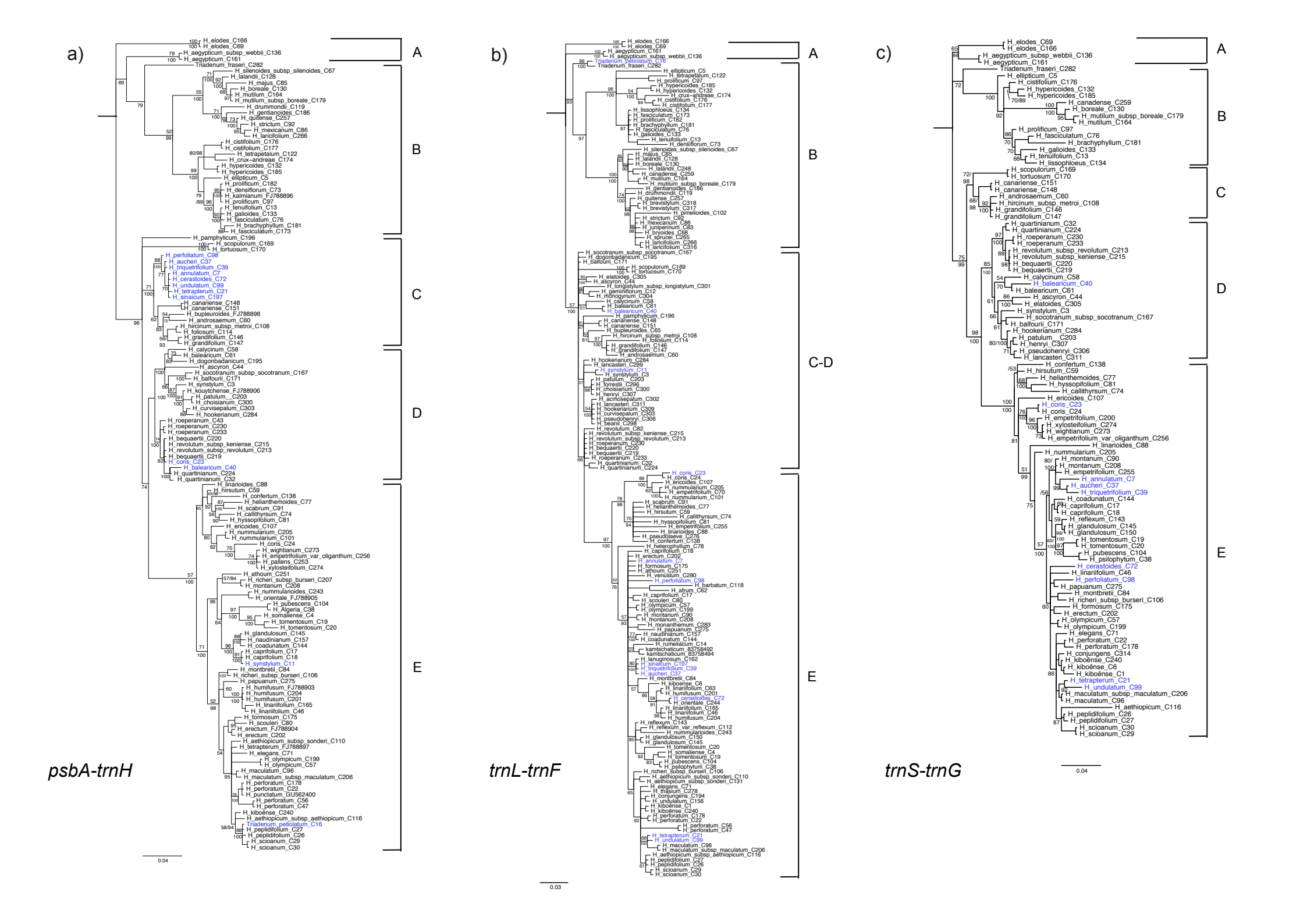
SI Figure 4. **Phylogenetic relationships in *Hypericum* inferred from the concatenated "All-specimens" plastid dataset (*trnF-trnL, trnS-trnG*).** Bayesian Majority-Rule consensus tree of *Hypericum* inferred from the combined chloroplast *trnF-trnL* and *trnS-trnG* markers. Bootstrap (bp) support values are indicated above branches for ML rearrangements (500 replicates in GARLI) and Bayesian posterior probabilities below branches (pp).

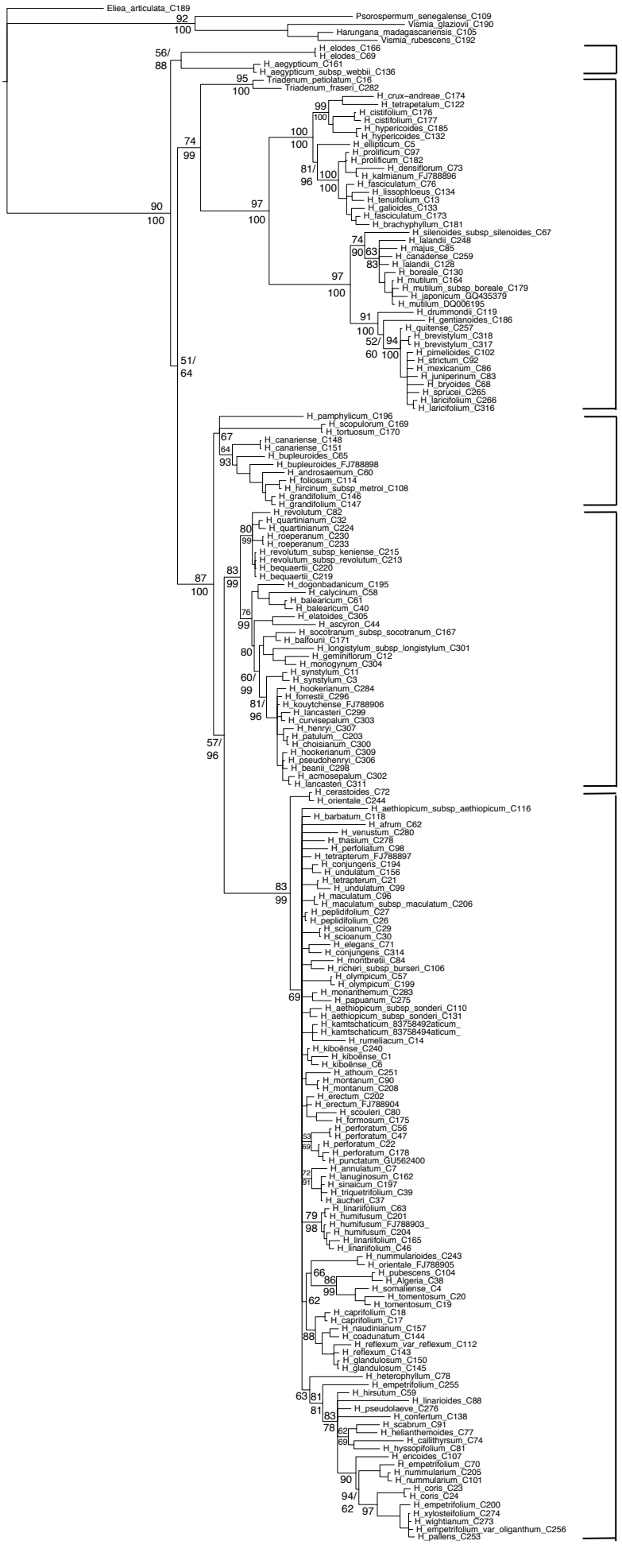

ITS
p distance
ML distance
0,5
0,45
0,4
0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7


psbA-trnH
p distance
ML distance
0,45
0,4
0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
0,1
0,2
0,3
0,4
0,5
0,6


trnL-trnF
p distance
ML distance
0,45
0,4
0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
0,05
0,1
0,15
0,2
0,25
0,3
0,35
0,4
0,45
0,5


trnS-trnG
p distance
ML distance
0,4
0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
0,05
0,1
0,15
0,2
0,25
0,3
0,35
0,4
0,45
0,5


a)
b)
c)
A
B
C
D
E
C-D
psbA-trnH
trnL-trnF
trnS-trnG
0.04
0.03
0.04


Eliea_articulata_C189
Psorospermum_senegalense_C109
Vismia_glaziovii_C190
Harungana_madagascariensis_C105
Vismia_rubescens_C192
H_elodes_C166
H_elodes_C69
H_aegypticum_C161
H_aegypticum_subsp_webbii_C136
Triadenum_petiolatum_C16
Triadenum_fraseri_C282
H_crux−andreae_C174
H_tetrapetalum_C122
H_cistifolium_C176
H_cistifolium_C177
H_hypericoides_C185
H_hypericoides_C132
H_elliptícum_C5
H_prolificum_C97
H_prolificum_C182
H_densiflorum_C73
H_kalmianum_FJ788896
H_fasciculatum_C76
H_lissophloeus_C134
H_tenuifolium_C13
H_galioides_C133
H_fasciculatum_C173
H_brachyphyllum_C181
H_silenoides_subsp_silenoides_C67
H_lalandii_C248
H_majus_C85
H_canadense_C259
H_lalandii_C128
H_boreale_C130
H_mutilum_C164
H_mutilum_subsp_boreale_C179
H_japonicum_GQ435379
H_mutilum_DQ006195
H_drummondii_C119
H_gentianoides_C186
H_quitense_C257
H_brevistylum_C318
H_brevistylum_C317
H_pimelioides_C102
H_strictum_C92
H_mexicanum_C86
H_juniperinum_C83
H_bryoides_C68
H_sprucei_C265
H_laricifolium_C266
H_laricifolium_C316
H_pamphylicum_C196
H_scopulorum_C169
H_tortuosum_C170
H_canariense_C148
H_canariense_C151
H_bupleuroides_C65
H_bupleuroides_FJ788898
H_androsaemum_C60
H_foliosum_C114
H_hircinum_subsp_metroi_C108
H_grandifolium_C146
H_grandifolium_C147
H_revolutum_C82
H_quartinianum_C32
H_quartinianum_C224
H_roeperanum_C230
H_roeperanum_C233
H_revolutum_subsp_keniense_C215
H_revolutum_subsp_revolutum_C213
H_bequaertii_C220
H_bequaertii_C219
H_dogonbadanicum_C195
H_calycinum_C58
H_balearicum_C61
H_balearicum_C40
H_elatoides_C305
H_ascyron_C44
H_socotranum_subsp_socotranum_C167
H_balfourii_C171
H_longistylum_subsp_longistylum_C301
H_geminiflorum_C12
H_monogynum_C304
H_synstylum_C11
H_synstylum_C3
H_hookerianum_C284
H_forrestii_C296
H_kouytchense_FJ788906
H_lancasteri_C299
H_curvisepalum_C303
H_henryi_C307
H_patulum_C203
H_choisianum_C300
H_hookerianum_C309
H_pseudohenryi_C306
H_beanii_C298
H_acmosepalum_C302
H_lancasteri_C311
H_cerastoides_C72
H_orientale_C244
H_aethiopicum_subsp_aethiopicum_C116
H_barbatum_C118
H_afrum_C62
H_venustum_C280
H_thasium_C278
H_perfoliatum_C98
H_tetrapterum_FJ788897
H_conjungens_C194
H_undulatum_C156
H_tetrapterum_C21
H_undulatum_C99
H_maculatum_C96
H_maculatum_subsp_maculatum_C206
H_peplidifolium_C27
H_peplidifolium_C26
H_scioanum_C29
H_scioanum_C30
H_elegans_C71
H_conjungens_C314
H_moritzbretii_C84
H_richeri_subsp_burseri_C106
H_olympicum_C57
H_olympicum_C199
H_monanthemum_C283
H_papuanum_C275
H_aethiopicum_subsp_sonderi_C110
H_aethiopicum_subsp_sonderi_C131
H_kamtschaticum_83758492aticum_
H_kamtschaticum_83758494aticum_
H_rumeliacum_C14
H_kiboënse_C240
H_kiboënse_C1
H_kiboënse_C6
H_athoum_C251
H_montanum_C90
H_montanum_C208
H_erectum_C202
H_erectum_FJ788904
H_scouleri_C80
H_formosum_C175
H_perforatum_C56
H_perforatum_C47
H_perforatum_C22
H_perforatum_C178
H_punctatum_GU562400
H_annulatum_C7
H_lanuginosum_C162
H_sinaicum_C197
H_triquetrifolium_C39
H_aucheri_C37
H_linariifolium_C63
H_humifusum_C201
H_humifusum_FJ788903_
H_humifusum_C204
H_linariifolium_C165
H_linariifolium_C46
H_nummularioides_C243
H_orientale_FJ788905
H_pubescens_C104
H_Algeria_C38
H_somaliense_C4
H_tomentosum_C20
H_tomentosum_C19
H_caprifolium_C18
H_caprifolium_C17
H_naudinianum_C157
H_coadunatum_C144
H_reflexum_var_reflexum_C112
H_reflexum_C143
H_glandulosum_C150
H_glandulosum_C145
H_heterophyllum_C78
H_empetrifolium_C255
H_hirsutum_C59
H_linarioides_C88
H_pseudolaeve_C276
H_confertum_C138
H_scabrum_C91
H_helianthemoides_C77
H_callithyrsum_C74
H_hyssopifolium_C81
H_ericoides_C107
H_empetrifolium_C70
H_nummularium_C205
H_nummularium_C101
H_coris_C23
H_coris_C24
H_empetrifolium_C200
H_xylosteifolium_C274
H_wightianum_C273
H_empetrifolium_var_oliganthum_C256
H_pallens_C253
A
B
C
D
E
0.03

SI Fig. 3


Eliea_articulata_C189
Psorospermum_senegalense_C109
Vismia_glaziovii_C190
Harungana_madagascarensis_C105
H_elodes_C69
H_elodes_C166
H_aegypticum_subsp_webbii_C136
H_aegypticum_C161
Triadenum_fraseri_C282
Triadenum_petiolatum_C16
H_ellipticum_C5
H_tetrapetalum_C122
H_hypericoides_C132
H_hypericoides_C185
H_crux-andreae_C174
H_cistifolium_C176
H_cistifolium_C177
H_prolificum_C97
H_galioides_C133
H_fasciculatum_C76
H_brachyphyllum_C181
H_fasciculatum_C173
H_prolificum_C182
H_lissophloeus_C134
H_tenuifolium_C13
H_densiflorum_C73
H_silenoides_subsp_silenoides_C67
H_boreale_C130
H_canadense_C259
H_lalandii_C248
H_majus_C85
H_lalandii_C128
H_mutilum_subsp_boreale_C179
H_mutilum_C164
H_gentianoides_C186
H_drummondii_C119
H_quitense_C257
H_brevistylum_C318
H_brevistylum_C317
H_pimelioides_C102
H_stictum_C92
H_mexicanum_C86
H_juniperinum_C83
H_bryoides_C68
H_sprucei_C265
H_laricifolium_C266
H_laricifolium_C316
H_scopulorum_C169
H_tortuosum_C170
H_pamphylicum_C196
H_canariense_C151
H_canariense_C148
H_bupleuroides_C85
H_androsaemum_C60
H_hircinum_subsp_metroi_C108
H_grandifolium_C146
H_grandifolium_C147
H_foliosum_C114
H_dogonbadanicum_C195
H_calycinum_C58
H_balearicum_C40
H_balearicum_C61
H_quartinianum_C32
H_quartinianum_C224
H_revolutum_C82
H_roeperanum_C230
H_roeperanum_C233
H_revolutum_subsp_revolutum_C213
H_revolutum_subsp_keniense_C215
H_bequaertii_C220
H_bequaertii_C219
H_ascyron_C44
H_elatoides_C305
H_socotranum_subsp_socotranum_C167
H_balfourii_C171
H_longistylum_subsp_longistylum_C301
H_geminiflorum_C12
H_monogynum_C304
H_synstylum_C3
H_synstylum_C11
H_patulum_C203
H_henryi_C307
H_hookerianum_C284
H_torrestii_C296
H_choisianum_C300
H_lancasteri_C299
H_pseudohenryi_C306
H_lancasteri_C311
H_acmosepalum_C302
H_hookerianum_C309
H_curvisepalum_C303
H_beanii_C298
H_confertum_C138
H_hirsutum_C59
H_linarioides_C88
H_pseudolaeve_C276
H_helianthemoides_C77
H_hyssopifolium_C81
H_callithyrsum_C74
H_scabrum_C91
H_ericoides_C107
H_nummularium_C205
H_empetrifolium_C70
H_nummularium_C101
H_coris_C23
H_coris_C24
H_xylosteifolium_C274
H_wightianum_C273
H_empetrifolium_var_oliganthum_C256
H_empetrifolium_C200
H_heterophyllum_C78
H_perfoliatum_C98
H_empetrifolium_C255
H_barbatum_C118
H_afrum_C62
H_athoum_C251
H_venustum_C280
H_montanum_C90
H_montanum_C208
H_formosum_C175
H_erectum_C202
H_papuanum_C275
H_monanthemum_C283
H_olympicum_C57
H_olympicum_C199
H_scouleri_C80
H_cerastioides_C72
H_orientale_C244
kamtschaticum_83758492
kamtschaticum_83758494
H_rumeliacum_C14
H_annulatum_C7
H_aucheri_C37
H_triquetrifolium_C39
H_lanuginosum_C162
H_sinaicum_C197
H_montbretii_C84
H_linariifolium_C63
H_humifusum_C201
H_linariifolium_C46
H_linariifolium_C165
H_humifusum_C204
H_caprifolium_C17
H_caprifolium_C18
H_coadunatum_C144
H_naudinianum_C157
H_tomentosum_C19
H_tomentosum_C20
H_pubescens_C104
H_Algeria_C38
H_reflexum_C143
H_nummularioides_C243
H_glandulosum_C145
H_glandulosum_C150
H_somaliense_C4
H_reflexum_var_reflexum_C112
H_richeri_subsp_burseri_C106
H_aethiopicum_subsp_sonderi_C110
H_aethiopicum_subsp_sonderi_C131
H_aethiopicum_subsp_aethiopicum_C116
H_conjungens_C314
H_elegans_C71
H_perforatum_C56
H_perforatum_C47
H_thasium_C278
H_peplidifolium_C26
H_peplidifolium_C27
H_perforatum_C22
H_perforatum_C178
H_scioanum_C30
H_scioanum_C29
H_conjungens_C194
H_undulatum_C156
H_kiboense_C240
H_kiboense_C6
H_kiboense_C1
H_tetrapterum_C21
H_undulatum_C99
H_maculatum_subsp_maculatum_C206
H_maculatum_C96
A
B
C
D
E
0.02

SI. Fig. 4

ANNEXES OF CHAPTER 4

SI TEXT

I. Climatic vegetation maps (Figure 1):

To draw the climatic-vegetation maps presented in Figure 1, we used paleobotanical evidence, as well as vegetation simulation models. The close-canopy tropical forest in the Southern Hemisphere is represented by light green colour. The northern vegetation belt in dark green colour represents alternatively the boreotropical forest (Fig. 1a), and its successors: the mixed-mesophytic forest (Fig. 1b,c) and the temperate forest (Fig. 1d). The boreal forest belt in the northern regions of Eurasia and North America (Fig. 1c) is represented by a darker shade of green.

a) The limits of the boreotropical forest during the Early Tertiary (Fig. 1a) were approximated based on the location of paleontological sites with boreotropical elements, following Wolfe (1975) and Tiffney (1985a,b). This extension is in accordance with the vegetation maps of Parrish (1982) and Fine & Ree (2006), although see Morley (2007, 2003) for a reduced distribution of the megathermal close canopy forest. Dashed lines show uncertainties in the distribution of the boreotropical forest.

b) The extension of the mixed-mesophytic forest during the Oligocene–Late Miocene (Fig. 1b) follows Wolfe (1975), Tiffney (1985a,b), Wen & Ickert-Bond (2009), and Beerling & Woodward (2001). It extends in the south into North of Africa (Wen & Ickert-Bond, 2009) and South East Asia (Wolfe, 1975; Tiffney, 1985a,b; Wen, 1999; Xiang et al., 2000). The paleoflora of Egypt during the Eocene/Oligocene was similar to the floras of Europe (Tiffney (1985b), and this was also the case for the mesothermal sclerophyllous vegetation in Taiwan (Wolfe, 1975). During the Early Tertiary, continental South East Asia (South China, Indochina, northwest Thailand, and West Malaysia) were covered by the same boreotropical flora as Japan and China (Sanmartín et al. 2001). The northern limit of the mixed-mesophytic forest is justified by the presence of paleontological sites in northern America at 45 degrees of actual latitude (Wolfe, 1975).

c) The extension of the temperate and boreal forest during the Pliocene is based on the climatic simulations of Salzmann et al. (2008), as well as on available paleobotanical evidence (Willis & McElwain, 2002; Wolfe, 1975, 1971; Tiffney, 1985a,b; Wen, 1999; and Schneck et al. 2012). Information about the vegetation of Australia across time periods comes from Morley (2003), Kemp (1978), and Frakes & Kemp (1972).

d) Position of the southern refuges for temperate forest taxa during the Last Glacial Maximum comes from Fine & Ree (2006) and Woillez et al. (2011). Milne (2006), however, excluded Central America and the eastern highlands of Mexico as potential refuges for the distribution of the temperate forest.

II. *Hypericum* fossil record

The medical plant genus *Hypericum* (St John's wort) has a rich fossil record that expands from the Late Eocene to the present (SI Table 3, Meseguer & Sanmartín, 2012). This record comprises some macrofossils (leaves), but mostly consists of microfossils remains, seeds and pollen. Among them, seeds are particularly distinctive in *Hypericum* and exhibit different sculpture patterns in the exotesta – ranging from reticulate, scalariform or papillose to smooth – which can characterize entire taxonomical sections (Robson, 1981; Meseguer & Sanmartín, 2012).

The oldest fossil remains of the genus (Meseguer & Sanmartín, 2012), the Late Eocene fossil seeds of *H. antiquum*, have been found in the Eastern Palearctic region, West Siberia (Arbuzova, 2005). The age and geographic distribution of this fossil were used to constrain estimates of divergence times and ancestral areas for crown-node *Hypericum* in the biogeographic analysis (see Material and Methods). The assignment of this fossil to the ancestor of all living *Hypericum* was based on the fact that it does not exhibit any particular feature (apomorphy) distinctive from other extant species. In fact, *H. antiquum* seeds have a characteristic ribbed (scalariform) sculpting pattern of the testa that is present in extant species of the basal section *Elodes* (clade A), the New World sister sections *Brathys* and *Trigynobrathys* (clade B), and the Old World section *Drosocarpium* (clade E) (Meseguer & Sanmartín, 2012). At the same time, this remain cannot be confidently assigned to the crown-node of any of these clades – it does not have any synapomorphy characteristic of a particular section within *Hypericum* (Meseguer & Sanmartin, 2012), so following Magallón & Sanderson (2001), we placed this fossil in the crown node of *Hypericum*, which is the most recent common ancestor of all sections with a scalariform testa sculpture (Meseguer et al., 2013).

The next oldest remain belongs to the fossil species *H. septestum,* and was also found in the Eastern Palearctic region in Lower Oligocene deposits (Tomsk, Russia; Arbuzova 2005). This species became a common element in the European fossil record of the genus from the

Late Oligocene to the Pliocene (Meseguer & Sanmartín, 2012). Seeds of *H. septestum* differ from the older *H. antiquum* remains in possessing reticulate patterns in the seed testa. This character state is by far the most abundant within the genus and appears scattered along the phylogeny (Meseguer et al., 2013). Therefore, the phylogenetic assignation of the fossil to a particular clade is very difficult. Meseguer et al. (2013) inferred that the reticulate seed testa is the ancestral state of the genus and of all major clades in the phylogeny. Scalariform patterns evolved in parallel in many unrelated clades, while a papillose testa is typical of the *Hirtella* group (clade D). However, there is also the possibility that the high extinction rates in the Eastern Palearctic that have affected the biogeographic reconstruction (see main text) may have obscured the inference of ancestral states and that the scalariform testa present in *H. antiquum* is indeed the ancestral state of *Hypericum* (Meseguer et al. 2013).

Although it varies across some species and clades, Meseguer et al. (2013) showed that the character of the testa pattern is not constrained phylogenetically and exhibits high levels of homoplasy, which prevents confident assignment of seed remains to phylogenetic clades. This applies to the other Oligocene to Pliocene fossils. Some *Hypericum* remains have been attributed to extant species, like *H. perforatum fossilis* or *H. androsaemum fossilis* (Mai & Walther, 1988), but we think this assignation needs to be further investigated. To our knowledge, there is no study for the whole genus providing diagnostic characters to identify *Hypericum* seeds at the species level. The sculpting pattern of the endotesta has been suggested as been specific for individual taxa (Ramos Núñez, 1983), but so far, this trait has only been described in some Iberian species and it is difficult to observe, i.e., the exotesta needs to be removed in order to examine the inner epidermal layer.

III. Ecological Niche Modelling (ENM) analyses

3.1. Species distribution data and climatic scenarios

Species occurrence data for extant *Hypericum* species were obtained from taxonomic monographies (Robson, 1981-onwards), herbarium collections, new-targeted fieldwork, and online databases (GBIF, http://www.gbif.org/), the latter collated to exclude ambiguous citations. In total, 1033 occurrences were included. Fossil distribution data (143 records, Table S3) was obtained from the literature (Meseguer & Sanmartín 2012) and online resources (Paleobiology Database, http://paleodb.org/). After excluding Pleistocene records, the data set comprises 68 greenhouse (19 pixels) and 65 coldhouse (18 pixels) occurrences.

Fossil records were grouped into two time slices to ensure an adequate number of occurrences to model ancestral species data, "greenhouse" lineages - covering the Eocene, Oligocene, and Early Miocene periods (68 occurrences, 18 pixels) - and "coldhouse" fossil lineages (65 occurrences, 19 pixels), spanning the Late Miocene and Pliocene (Pleistocene fossil records were excluded) (Table S3). The boundary between greenhouse and coldhouse time slices was set at the Late Miocene cooling event, after the dramatic drop in temperature following the Middle Miocene Climatic Optimum (ca. 11.5 Ma; Zachos et al. 2008). This is a period when important changes on vegetation occurred (Schneck et al., 2012), such as the replacement of the mixed-mesophytic forest by a boreal coniferous forest across the northern regions of Eurasia and North America. Fossil paleocoordinates were calculated with the PointTracker application of the Paleomap software (Scotese, 2010).

For past climate scenarios, we used a suite of six different global climate simulations by the Hadley Centre coupled ocean-atmosphere general circulation model (HadCM3L): a) an Early Eocene climate simulation with representation of early Eocene (55Ma) paleogeography (continental coastline and icesheet configuration) and 1120 p.p.m CO_2 concentration to represent peak Eocene warmth (but not the transient spike in warming at the Paleocene-Eocene thermal maximum) (Beerling et al., 2011); b) the same HadCM3L model set-up but with a lower CO_2 concentration of 560 p.p.m, to represent the drop in temperature at the Terminal Eocene Event. Details of the HadCM3L model set-up for these Eocene simulations are given in Beerling et al. 2011; c) a 400 p.p.m CO_2 Late Miocene simulation representing the early warm period of the Miocene (Mid Miocene Climatic Optimum, 15 Ma); d) Miocene model set-up with 280 p.p.m CO2 concentration, representing the cold and dry conditions important after the Late Miocene cooling event. Details of the HadCM3L set-up for these Miocene simulations are given in Beerling et al. (2012) and Bradshaw et al. (2012), e) a 560 ppm CO2 Pliocene simulation (Beerling et al., 2009) of the conditions at the Mid-Pliocene warming event (3.6–2.6 Ma); and f) a simulation of the Preindustrial World with 280 ppm CO2 to provide a baseline HadCM3L climate before the industrial revolution from Beerling et al. (2012).

3.2 Selection of climatic variables and modelling techniques

To discriminate the most relevant climatic variables able to explain *Hypericum* occurrences we used extant occurrences and temperature and precipitation data for current conditions

(≈1950–2000) extracted from WorldClim (Hijmans et al., 2005). We used mean monthly data at a resolution of 10 arc-minutes (around 20 x 20 km) deriving nine climatic variables representing important vegetation predictors capable of being equally estimated for past scenarios: annual precipitation, annual variation in precipitation, aridity, continentality, maximum monthly precipitation, maximum monthly temperature, mean annual temperature, minimum monthly precipitation and minimum monthly temperature. Bioclimatic indices as aridity and continentality were calculated following the formulas provided by Valencia-Barrera et al. (2002).

We estimated the minimum set of climatic variables needed to explain *Hypericum* occurrences using an ecological-niche factor analysis (ENFA; Hirzel et al., 2002; Calenge & Basille, 2008) with the terrestrial world as the background area. This procedure compares the climatic data of presence localities against the climatic conditions found throughout the study area, thereby computing uncorrelated factors that can explain both species marginality (the distance between the species optimum and the average climatic conditions in the study area) and specialization (the ratio of the ecological variance in the climate of the study area to that associated with the focal species). Factors were retained based on their eigenvalues relative to a broken-stick distribution (Hirzel et al., 2002). Climatic variables selected as predictors were those showing the highest correlation values (factor scores > 0.30) with the retained ENFA factors; i.e. those able to better discriminate the climatic conditions in the presence localities against world conditions. Including the former mentioned nine climatic variables we obtain two factors that explain 91.3% of total variability and 82.6% of specialization. *Hypericum* marginality value is 0.54, so optimum climatic conditions were relatively near from the average conditions available in the world, and specialization value is 2.46 (i.e. the *Hypericum* range of habitable conditions is more than two times narrower than those available in the whole terrestrial Earth). The first ENFA factor (Fig. S5) was positively related with annual precipitation (factor score, *fs* = 0.528), annual variation in precipitation (*fs* = 0.253), and maximum and minimum monthly precipitation variables (*fs* =0.446 and 0.448, respectively) but negatively with continentality (*fs* = -0.314). The second factor (Fig. S5) was negatively correlated with aridity (*fs* = -0.644) and minimum monthly temperature (*fs* = -0.659). Thus, we may consider that the first ENFA factor represents “tropicality”, while the second factor represents “aridity”. These seven climatic variables were selected for further analyses.

Instead of using modelling techniques based on statistical fitting of data or machine learning procedures capable of capturing complex spatial patterns, we use simple methods derived from geometrical or set theory than only use the information on observed presences. Presence data is the only reliable source of information capable of provide information on the climatic conditions for which the species may have a positive net rate of demographic growth. Both statistical and machine learning techniques require pseudoabsences generally selected at random from the whole distribution area, a risked procedure in the case of the incomplete fossil data because it is unable to reflect occupancy probabilities but the own density of observations used in the analysis (Aarts et al. 2012). Thus, we use a generalized intersection procedure by estimating extreme maximum and minimum values in the observed localities for the seven climatic variables previously selected, to subsequently transfer these conditions to the geographical space in order to represent the regions with similar conditions to those where the taxa occurred (potential distribution). The so obtained binary representation was transformed to a continuous one by calculating the scale-invariant Mahalabobis Distance (MD) from the average climatic conditions representing lineage's hypothetical climatic optimum (Varela et al., 2011).

3.3. Climatic segregation among extant clades

We used the values of the two ENFA factors obtained above to examine climatic niche differences among extant major clades in Meseguer et al. (2013)'s phylogeny. We assessed niche segregation across phylogenetic clades using the values of these two factors in all recorded localities where each species belonging to a clade are found, and tested for significant differences among clades by using a non-parametric Kruskal-Wallis by rank test and post-hoc comparisons between all pairs groups.

3.4 Fossil and present ENM projections

We projected the climatic distance from i) present-day optimum based on extant data, ii) "coldhouse" optimum based on fossils, and iii) "greenhouse" optimum based on fossils, onto the six past climatic simulations to build a continuous geographic representation model, using the Mahalanobis distance to obtain a continuous measure of climatic favourability. These models allowed us to identify areas in the underlying paleo-continental reconstruction (Markwick 2007) with suitable climate for *Hypericum* during the relevant paleotime frame

(Figs S1-S2). To assess the degree of niche stability versus niche evolution along the history of *Hypericum*, we compared the geographic and climatic representations for greenhouse lineages when projected onto the more recent Late Miocene and Pliocene climate layers with the projection of the coldhouse climate optimum onto the Eocene and Early Miocene layers. In all these cases both current and paleocoordinates are contracted to a resolution similar to the used paleomaps (2.50 x 3.75 degrees). We use similar Kruskal-Wallis by rank test and post-hoc comparisons to estimate the temporal variation in the *Hypericum* favourable climatic conditions both for the whole terrestrial Earth pixels and for the pixels with fossil observations (Fig S3). To test the sensitivity of our analyses to the boundary selected between both time slices, especially in relation to the very variable Miocene period (Zachosetal2008}), we performed a second analysis in which we removed all Miocene records, leaving only the Eocene-Oligocene and Pliocene fossil sites (Fig S3). Since greenhouse and coldhouse fossil lineages span several climatic layers, we estimated their climatic optimum with respect to the earliest and latest climatic simulations, respectively, i.e., Late Eocene (greenhouse) and Pliocene (coldhouse).

IV. Biogeographic analysis

4.1 Area selection

We split the world into seven operational areas. Delimitation was based on patterns of endemism within *Hypericum*, i.e., shared distribution patterns among species, but also on the palaeogeographic history of the continents with the aim to maximize congruence with other continental-scale biogeographic studies (Sanmartín et al. 2001; Donoghue & Smith, 2004). Several areas coincide with the floristic regions of Takhtajan (1986), which was based on patterns of vegetation endemism, and therefore likely reflect geographical and eco-physiographical (climatic) features. Western Palaearctic region ("A"): Eurasia west of the Ural Mountains (former Turgai Sea); Eastern Palaearctic ("B"): non-tropical Asia east of the Ural Mountains; Irano-Turanian-Himalayan ("C"): the region formed by the collision of the Indian and Arabian plates with Eurasia, extending from the Zagros Mountains and the Caucasus in the west through the Iranian plateau, the Hindu Kush, Tien Shan and Kunlun Shan Mountains to the Altay Mountains in the northeast, and the Tibetan Plateau and the Himalayan Mountain range in the southeast; Nearctic ("D"): North America and Central America down to the Mexican lowlands (i.e., excluding tropical Mexican distributions); the

Caribbean region was included here because there are very few endemics and most widespread species are shared with eastern North America; Oceania ("E"): comprising the tropical regions of southwest and southern China and eastern parts of Bangladesh, Burma, Thailand, Indochina, the Malay Peninsula and South East islands of Borneo, Sumatra, Java, and the Inner Banda Arc, the Polynesian region, and Hawaii, as well as the Indian subcontinent (all Indostan excluding the tropical foothills of Himalayas in the north) and the former Gondwanan landmasses of Australia, New Guinea, New Caledonia, and New Zealand; Africa ("F"): excluding the region north of the Saharan belt, which is considered as part of the Western Palearctic region; Neotropics ("G"): South America and tropical Central America up to central Mexico.

4.2 Phylogenetic and lineage divergence estimation

To reconstruct ancestral geographic ranges and biogeographic events in *Hypericum*, we used the molecular dated phylogeny of Meseguer et al. (2013), which included 114 species representing ca. 25% of the described Hypericeae diversity (500 species) and 33 out of 36 recognized morphological sections (Robson, 2012). Five outgroup taxa representing all Hypericaceae tribes were included: *Harungana* and *Vismia* for tropical tribe Vismieae, sister to Hypericeae, and *Eliea* representing tropical tribe Cratoxyleae (Rufhel et al., 2011); genus *Triadenum*, which is nested within *Hypericum* in most molecular studies (Rufhel et al. 2011; Nürk et al., 2012; Meseguer et al., 2013), was also included. The final dataset comprises 3032 characters, which are 4 times more characters than similar studies at infrageneric level on *Hypericum* (Nürk et al., 2012; Park & Kim, 2004; Crocket et al., 2004). The sample covered almost all geographical regions where *Hypericum* is distributed and most of the morphological variation described in the group. The phylogenetic hypothesis was derived from three chloroplast markers (*trnL-trnF, trnS-trnG, psbA-trnH*) using Bayesian inference methods implemented in the software Mr. Bayes v3.2 (Ronquist et al. 2012). Meseguer et al. (2013) analysed different datasets varying in the amount of missing data to evaluate its effects on phylogenetic inference. For this study, we used the "two-chloroplast dataset", which only includes those specimens represented in at least two chloroplast markers. This dataset recovers the same overall topology and groupings as the complete ("All-specimens") data set (Meseguer et al., 2013), but with higher resolution and clade support values, and was the one used for Bayesian inference of morphology and range evolution by these authors.

Absolute divergence ages were estimated by analysing this dataset under a Bayesian relaxed clock model in BEAST v.1.6. (Drummond & Rambaut, 2007), using a by-gene partitioned dataset, Yule tree prior, and lognormal uncorrelated relaxed clock model (Meseguer et al., 2013). Two calibration points were used: the Late Eocene fossil *H. antiquum* to constrain the crown node of Hypericeae, and a secondary calibration point for the root of the tree (crown-node Hypericaceae) from the calibrated clusioid clade phylogeny published by Ruhfel (2011) (see Meseguer et al., 2013 for more details).

4.3. Ancestral range inference

Biogeographic analyses were conducted with the parametric likelihood method "Dispersal-Extinction-Cladogenesis" implemented in Lagrange (Ree et al. 2005; Ree & Smith, 2008) using the fast C++ version (Smith, 2009). We set up four different analyses (M1 to M4) in order of increasing complexity to analyse the influence of fossils and ENM reconstructions (past and present environmental tolerances) in biogeographic inference. Analyses were run on the maximum clade credibility tree from the MCMC BEAST analysis of Meseguer et al., 2013 (Fig. 2d). Additionally, we run the integrative M4 model over 1000 trees sampled from the BEAST posterior distribution to account for uncertainty in topology and branch length estimates. Changes in ancestral area inference resulting from model and phylogenetic uncertainty are presented in Table S1 and S2, respectively.

a) **Model M1** ("Unconstrained", Fig. S6): We used only present data from current distributions.

b) **Model M2** ("Stratified", Fig. S7): We used present occurrences but also incorporated a palaeogeographic model to reflect changes in continental connectivity over time. The evolution of *Hypericum* during the last 55 million years has been accompanied by dramatic changes in continental configuration: tectonic plates moved apart and collided with each other (e.g., Africa and Eurasia), but also the sea level changed producing the emergence and sinking of land corridors (e.g., Beringia), whereas climatic oscillations created ecological barriers to dispersal, e.g. the ice sheet covering part of the Northern Hemisphere in the Pleistocene. These changes probably had an effect on the probability of *Hypericum* lineages to disperse between areas over time.

To incorporate this information into our analysis, we stratified the phylogeny into four time slices (TS), with boundaries selected to reflect major tectonic or climatic events: **TSI**

(65 to 35 Ma): this period is characterized by warm climates at northern latitudes and the geographical proximity of northern landmasses, allowing the spread of a continuous forest belt, the *boreotropical* forest, across the Holarctic (Wolfe, 1975). It ends with the Terminal Eocene Event (TEE). **TSII** (35 to 10 Ma): this period is marked by the major Antarctic glaciations and the Terminal Eocene Event (TEE; Zachos et al., 2008), a global climatic deterioration that produced a diversity decline at the beginning of the Oligocene (Morley, 2007). During this period the boreotropical forest was replaced by the mixed-mesophytic forest, a mix of deciduous and conifer vegetation (Tiffney 1985a,b; Wolfe, 1975) that persisted until the Late Miocene (Wen, 1999; Xiang et al., 2000). **TSIII** (10 to 3.5 Ma) begins with the Late Miocene cooling event (LCM, ca. 10-8 Ma), a period when important changes in vegetation occurred (Schneck, 2012): a boreal forest expanded across the northern regions of Eurasia and North America across Beringia, whereas the mixed-mesophytic elements retreated to temperate regions in the south. **TSIV** (3.5-0 Ma) spans the period from the opening of the Bering Strait (3.5 Ma), which interrupted biotic connections between Palearctic and Nearctic regions for temperate elements to the glacial cycles that characterized the end of the Pliocene and the Pleistocene period.

Dispersal rates in the biogeographic model (Q transition matrix, Buerki et al., 2011) were scaled to reflect the changing continental configuration for each TS as follows: a) dispersal rates between areas sharing an edge (adjacent plates) or connected by a land corridor were not downscaled, i.e., assigned a scalar of 1, to reflect the facility of lineages to migrate across this type of connection; for example, the NALB between the Western Palearctic and North America in the Early Eocene. b) Dispersal rates between neighbouring areas not connected by a land bridge, such as the Western and Eastern Palearctic landmasses before the closing of the Turgai Strait, were downscaled by a factor of 0.5 to reflect the reduced probability of movement. c) Dispersal between non-neighbouring areas or areas separated by large ocean barriers was downscaled by a factor of 0.1; this also applies when dispersal implies an intermediate area, for example, the possibility of dispersal between the Eastern Palearctic and Africa through the Indian Plate in the Eocene. Besides paleogeographic barriers, other abiotic factors such as marine or wind currents (e.g., the Antarctic Cicum-Polar Current), or climatic "barriers" such as ice sheets, were considered in modelling dispersal rates over time.

Paleostratigraphic model:

North America and **Western Palearctic (D and A)**: During the Paleogene (55-35 Ma), North America and Europe became connected by the North Atlantic Land Bridge (NLAB). The first NLAB, the Thulean Bridge, connected the British Isles with the Hudson Bay in North America and allowed movement of warm-temperate boreotropical elements between 55 and 50 Ma, when it foundered. A later bridge, the De Geer Bridge, lasted until ca. 35 Ma, but, because of its northern position, was more appropriate for cold-temperate plants (Tiffney, 1985ab; Sanmartín et al., 2001). Our model reflects this connection as a dispersal probability of 1 for TSI, which decreases to 0.5 after the break up of the NLAB connection in the Late Eocene (TSII, 35 Ma). Finally, a low probability of 0.1 in the last two time slices, TSIII and TSIV, reflects the widening of the North Atlantic Ocean and increasing geographical distance between Nearctic and Palearctic regions after 20 Ma.

North America and **Eastern Asia (D and B)**: Land dispersal between North America and Asia was possible across the Beringian Bridge throughout the Early-Mid Tertiary (Tiffney, 1985a,b; Wen, 1999; Donoghue et al., 2001; Sanmartín et al., 2001), as reflected in our model (dispersal probability of 1 in TSI and TSII). As the climate became increasingly cooler from the Late Miocene onwards, the Beringian Bridge became covered by a boreal taiga forest that extended across northern Eurasia and northern North America (Sanmartín et al., 2001). This might have reduced the probability of dispersal between these two regions for temperate taxa (scaling factor 0.5 in TSIII). Finally, the opening of the Bering Strait 3.5 Ma broke all final connections between the Nearctic and Palearctic regions for boreo-temperate elements (TSIV, scaling factor 0.1), although tundra-arctic elements could still migrate across Beringia during the Pleistocene glaciations (Sanmartín et al., 2001).

North America and **South America (D and G)** were intermittently connected during the Tertiary, first by the Late Eocene Proto-Greater Antilles land bridge, and later by GAARlandia (Greater-Antilles Aves-Ridge) during the Eocene-Oligocene (33-35 Ma) boundary (Briggs, 1994). This is reflected in our palaeostratigraphic model by a dispersal scaling factor of 0.5 in TSI and TSII. This land bridge sank and the connection was interrupted or reduced until the Late Miocene (ca. 10 Ma), when new mountain uplift in Central America and South America created a new migration route: the Late Miocene Central America Land Bridge and northern Andes orogeny (Antonelli et al., 2009). Finally, after the uplift of the Panama Isthmus (3.5 Ma), land dispersal has been possible across the New World (scaling factor of 1 in TSIII and TSIV)**.**

Western Palearctic and Eastern Palearctic (A and B) were separated during the Early Tertiary by the Turgai Strait, although connections along the coasts of the Tethys Seaway were possible (Tiffney, 1985b; Sanmartín et al., 2001). Thus, dispersal between these two regions was scaled by a factor of 0.5 in TSI. The Turgai Strait closed at the Eocene-Oligocene boundary (30 Ma) with the formation of the Ural Mountains, and from this moment on direct land dispersal was allowed in our biogeographic model (scaling factor of 1 from TSII to TSIV).

The **Irano-Turanian-Himalayan (C)** region is formed by the newly uplifted landmasses and mountain ranges that resulted from the successive collision against Eurasia of the Indian plate in the Eocene and the Arabian plate in the Middle Miocene (Sanmartín, 2003; Wang et al., 2009). Before the Tertiary, this area formed the eastern arm of the Tethys Seaway, and therefore dispersal to this region by the other landmasses was disallowed until TSIII, i.e., by assigning it a probability of 0**.**

Africa and Eurasia (**F and A-B**): At the beginning of the Paleocene, 60 Ma, the north-eastward drift of the African plate brought this continent into contact with the western half of Eurasia, the Iberian Peninsula. The Gibraltar Strait was intermittently opened and closed during the Tertiary (Sanmartín, 2003; Meulenkampt & Sissingh 2003), but the African and Western Palearctic plates have always been geographically close and this proximity is reflected in our model by allowing dispersal between these two areas across all time slices (scale factor of 1 from TSI to TSIV). In contrast, direct land dispersal between **Africa** and the **Eastern Palearctic** only became possible in the Mid Miocene (16 Ma, TSII), when the Arabian plate collided with Eurasia along the Anatolian fault zone, closing the eastern arm of the Tethys Sea (Sanmartín, 2003; although alternative reconstructions suggest an earlier collision, 30 Ma, Allen & Amstrong 2008). The Arabian bridge acted as a dispersal corridor between Africa and central-western Asia until the Late Miocene-Pliocene, when progressive aridification of the African continent led to the formation of the Saharan (7 Ma) and Arabian Deserts (3.5 Ma). These wilderlands probably constituted a barrier to dispersal, reflected in our model by downscaling dispersal to 0.5 in TSIII and IV.

South America and Australia (Oceania) (G and E): These two continents were connected as part of Eastern Gondwana until the opening of the Drake Passage between South America and Antarctica in the Early Oligocene (32 Ma). Therefore, a scaling factor of 1 was assigned to TSI to reflect the possibility of land dispersal. The opening of the Drake

Passage isolated these regions by opening new ocean barriers and initiated the Antarctic Circum-Polar Current, which brought about the first Antarctic glaciation (Morley, 2003, Sanmartín, 2005). This is reflected in our model by a low probability of dispersal from TSII to TSIV (scalar 0.1).

Eastern Palearctic and Australia (Oceania) (B and E): During the Early Tertiary, these two landmasses were separated by the Tethyan Gulf (scalar 0.1 in TSI). Following its separation from Antarctica in the Late Eocene-Early Oligocene, Australia began to drift rapidly towards Asia. The collision of the Australian Plate with the Pacific and Eurasian Plates resulted in the formation of the South West Pacific and South East Asian archipelagos (Metcalfe, 1998), which might have acted as a route of dispersal. This is reflected in our model by allowing dispersal between areas B and E, albeit with a low rate (0.5) in TSII. Dispersal, however, was freely allowed in TSIII and TSIV, after the final collision of the Australian plate with Eurasia.

Irano-Turanian-Himalayan and Eurasia (C and A-B): Movement between these two regions might have been possible after the onset of the Equatorial Ocean Currents in the Late Tertiary, involving dispersal along the coasts of India, which forms part of area E (Sanmartín & Ronquist, 2004). This is reflected in our model by a dispersal rate of 0.5 in TSIII and TSIV.

c) **Model M3** ("Fossil-stratified", Fig. S8):

We used present occurrences but also integrated information from fossil ranges and paleogeographic scenarios into the reconstruction. In particular, the crown node of *Hypericum* was constrained to include the Eastern Palearctic region ("B") as part of the ancestral range of this node. This is the area of appearance of the fossil, *H. antiquum*, which was used to calibrate the age of the crown-node in the phylogenetic dating analysis. In our approach, fossil constraints placed on nodes are encoded as ancestral presence in the area containing the geographic locality of the fossil. Lagrange then calculates the likelihood of the data conditional on that area being included in the ancestral range at the node in question (these scripts are available from R. Ree on request). The paleogeographic model in Model 2 was also used to constrain area connectivity through time.

d) **Model M4** ("Integrative", Fig. S9):

We integrated fossil constraints and fossil-based ENM predictions (Fig. S2) in the biogeographic analysis in two ways. First, by running Lagrange with the crown node of *Hypericum* constrained to include the Eastern Palearctic and Nearctic regions; these regions were predicted as potential distribution areas for *Hypericum* lineages during the Eocene-Oligocene by the Eocene 560 ppm projections (Fig. 2, Fig. S2). Second, we modified the paleogeographic dispersal matrix in Model 2 to reflect the existence of climatic/ecological barriers or transient dispersal corridors that connected regions, e.g., areas that are now outside the climatic envelope of the organism but were in the past within its ecological tolerance. Therefore, this new model reflects both the geographical and "ecological" connectivity between regions in *Hypericum*. We ran the analysis over the posterior sample of 1000 trees from BEAST to account for topological and temporal uncertainty.

Paleostratigraphic-ecological model: The above-defined probabilities of movement between areas were increased or decreased in accordance to the adequacy of the world at different time periods to the present and past tolerances of *Hypericum* species (Fig 2a). Areas were considered ecologically connected when there were regions/pixels within the ecological tolerance of the group in a given time period, i.e., with Mahalanobis distances lower than ten. Specifically:

TSII: The dispersal rate scalar between **Africa** and **Eastern Palearctic (F and B)** was increased to 0.5 because ENM models favoured this Miocene dispersal corridor for *Hypericum* (Fig 2a).

TSIII: The dispersal rate scalar between **Eastern Palearctic** and **Nearctic** was increased to 0.7. Several studies suggest that during the Late Miocene-Pliocene the Beringian corridor was only favourable for cold-adapted (boreal) taxa (Sanmartín et al., 2001), but our ENM models suggest that it was still within the environmental tolerance of *Hypericum* at least until the Pliocene (Fig. 2a). On the other hand, despite continuous geographic connection between the **Western Palearctic** and **Africa** after the Paleogene, the dispersal rate between these two areas was downscaled to 0.7 in this time slice, because our ENM models suggest that the aridification trend that affected North Africa from the Late Miocene onwards (7-5 Ma) probably created an ecological barrier for the dispersal of *Hypericum* lineages (Fig. 2a)

TSIV: Again, the ENM models predict a decrease in connectivity between **Africa** and **Western Palearctic** in the Pliocene (Fig. 2a), related to the aridification episode that gave

rise to the Saharan and Arabian deserts, so dispersal rates between these two regions was downscaled to 0.5 in this time slice.

BIBLIOGRAPHY

Aarts G, Fieberg J, Matthiopoulus J (2011) Comparative interpretation of count, presence-absence and point methods for species distribution models. *Methods Ecol Evol* 3: 177–187.

Allen MB, Armstrong HA (2008) Arabia-Eurasia collision and the forcing of mid-Cenozoic global cooling. *Palaeogeogr Palaeoclimatol Palaeoecol* 265: 52–58.

Antonelli A, Nylander JAA, Persson C, Sanmartín I (2009) Tracing the impact of the Andean uplift on Neotropical plant evolution. *Proc Natl Acad Sci USA* 106: 9749–9754

Arbuzova O (2005) *Hypericum* L. in: Iskopaemye tsvetkovye rastenija Rossii i sopredel'nyh gosudarstv [Fossil flowering plants of Russia and adjacent countries], eds Budantsev L (Izdatelstvo Nauka Leningradskoe otdnie, Leningrad), pp 1974-.

Beaumont LJ, Hughes L, Poulsen M (2005) Predicting species distributions: use of climatic parameters in BIOCLIM and its impact on predictions of species' current and future distributions. *Ecol Model* 186:250–269.

Beerling DJ et al. (2012) Ecosystem CO2 starvation and terrestrial silicate weathering: mechanisms and global-scale quantification during the late Miocene. *J Ecol* 100: 3141.

Beerling DJ, Foxa A, Stevenson DS, Valdes PJ (2011) Enhanced chemistry-climate feedbacks in past greenhouse worlds. *Proc Natl Acad Sci USA* 108: 9770–9775.

Beerling DJ, Royer DL (2011) Convergent Cenozoic CO2 history. *Nature* 4: 418–420

Beerling DJ, Woodward FI (2001) Vegetation and the terrestrial carbon cycle: modelling the first 400 million years. Cambridge University Press, Cambridge.

Beerling D, Berner RA, Mackenzie FT, Harfoot M, Pyle JA (2009) Methane and the CH_4-related greenhouse effect over the past 400 million years. *Am J Sci* 309:97–113.

Bradshaw CD et al. (2012) The relative roles of CO2 and palaeogeography in determining late Miocene climate: results from a terrestrial model-data comparison. *Clim Past* 8:1257–1285.

Briggs JC (1994) The genesis of Central America: biology versus geophysics. *Glob Ecol Biogeogr Lett* 4:169–172.

Buerki S, Forest F, Alvarez N, Nylander JAA, Arrigo N, Sanmartín I (2011) An evaluation of new parsimony-based versus parametric inference methods in biogeography: a case study using the globally distributed plant family Sapindaceae. *J Biogeogr* 38: 531–550.

Calenge C, Basille M (2008) A general framework for the statistical exploration of the ecological niche. *J Theor Biol* 252:674–685.

Crockett SL, Douglas AW, Scheffler BE, Khan IA (2004) Genetic profiling of *Hypericum* (St. John's Wort) species by nuclear ribosomal ITS sequence analysis. *Planta Med* 70:929–935.

Donoghue MJ, Smith SA (2004) Patterns in the assembly of temperate forests around the Northern Hemisphere. *Philos Trans R Soc Lond B Biological Sciences* 359:1633–1644.

Donoghue MJ, Bell CD, Li J (2001) Phylogenetic patterns in Northern Hemisphere plant geography. *Int J Plant Sci* 162(6): S41–S52.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *Bmc Evol Biol* 7:8.

Fine PVA, Ree RH (2006) Evidence for a time-integrated species–area effect on the latitudinal gradient in tree diversity. *Am Nat* 168:786–804.

Frakes LA, Kemp EM (1972) The influence of continental positions on early Tertiary climates. *Nature* 240: 97–100.

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. Jarvis, 2005. Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* 25:1965–1978.

Hirzel AH, Hausser J, Chessel D, Perrin N (2002) Ecological-niche factors analysis: how to compute habitat-suitability maps without absence data? *Ecology* 83: 2027–2036.

Kemp EM (1978) tertiary climatic evolution and vegetation history in the southeast Indian Ocean region. *Palaeogeogr Palaeoclimatol Palaeoecol* 24: 169–208

Magallon S, Sanderson MJ (2001) Absolute diversification rates in angiosperm clades. *Evolution* 55:1762–1780.

Mai DH, Walther H (1988) Die pliozänen Floren von Thüringen. *Quartärpaläontologie* 7: 55–295.

Markwick PJ (2007) The palaeogeographic and palaeoclimatic significance of climate proxies for data-model comparisons in Deep-Time Perspectives on Climate Change: Marrying the Signal From Computer Models and Biological Proxies, eds Williams M, Haywood AM, Gregory FJ, Schmidt DN (The Geological Society, London), pp 251–312.

Meseguer AS, Aldasoro JJ, Sanmartín I (2013) Bayesian inference of phylogeny, morphology and range evolution reveals a complex evolutionary history in St Johns wort (*Hypericum*). *Mol Phylogenet Evol* 67:379–403.

Meseguer AS, Sanmartín I (2012) Paleobiology of the genus *Hypericum* (Hypericaceae): A survey of the fossil record and its palaeogeographic implications. *Anales Jard Bot Madrid* 69:97–106.

Metcalfe I (1998) Paleozoic and Mesozoic geological evolution of the SE Asia region: multidisciplinary constraints and implications for biogeography in *Biogeography and geological evolution of SE Asia*, eds Halle R, Holloway JD (Backhuys, Leiden), pp 25–41.

Meulenkamp JE, Sissingh W (2003) Tertiary palaeogeography and tectonostratigraphic evolution of the Northern and Southern Peri-Tethys platforms and the intermediate domains of the African–Eurasian convergent plate boundary zone. *Palaeogeogr Palaeoclimatol Palaeoecol* 196:209–228.

Milne RI (2006) Northern hemisphere plant disjunctions: a window on Tertiary land bridges and climate change. *Ann Bot* 98, 465–472.

Morley RJ (2003) Interplate dispersal routes for megathermal angiosperms. *Perspectives Pl Ecol Evol Syst* 6: 5–20.

Morley RJ (2007) Cretaceous and Tertiary climate change and the past distribution ofmegathermal rainforests in: Tropical Rainforest Responses to Climatic Change, eds Bush MB, Flenley JR (Springer, Berlin), pp 1–31.

Nürk NM, Madriñan S, Carine MA, Chase MW, Blattner FR (2012) Molecular phylogenetics and morphological evolution of St. John's wort (Hypericum; Hypericaceae). *Mol Phylogenet Evol* 66: 1–16.

Park S, Kim K (2004) Molecular phylogeny of the genus *Hypericum* (Hypericaceae) from Korea and Japan: evidence from nuclear rDNA ITS sequence data. *J Plant Biol* 47:366–374.

Parrish JT, Ziegler AM, Scotese CR (1982) Rainfall patterns and the distribution of coals and evaporites in the Mesozoic and Cenozoic. *Palaeogeogr Palaeoclimatol Palaeoecol* 40:67–101.

Ramos Núñez A (1983) Estudio biosistemático del género *Hypericum* L. (Guttiferae) en la Península Ibérica e Islas Baleares. 1. Caracteres seminales. *Trabajo del Departamento de Botánica* 12: 45–62.

Ree RH, Moore BR, Webb CO, Donoghue MJ (2005) A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution* 59: 2299–2311.

Ree RH, Smith SA (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Syst Biol* 57:4–14.

Robson NKB (1981) Studies in the genus *Hypericum* L. (Guttiferae). 2. Characters of the genus. *Bull Brit Mus (Nat Hist) Bot* 8: 55–226.

Ronquist F et al. (2012) Mrbayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61:539–542.

Ruhfel BR et al. (2011) Phylogeny of the clusioid clade (Malpighiales): Evidence from the plastid and mitochondrial genomes. *Am J Bot* 98: 306–325.

Ruhfel BR (2011) Systematics and Biogeography of the Clusioid Clade (Malpighiales). Harvard University, Cambridge, Massachusetts.

Salzmann U, Haywood AM, Lunt DJ, Valdes PJ, Hill DJ (2008) A new global biome reconstruction and data-model comparison for the Middle Pliocene. *Global Ecol Biogeogr* 17: 432–447.

Sanmartín I, Enghoff H, Ronquist F (2001) Patterns of animal dispersal, vicariance and diversification in the Holarctic. *Biol J Linn Soc* 73:345–390.

Sanmartín I (2003) Dispersal vs. vicariance in the Mediterranean: historical biogeography of the Palearctic Pachydeminae (Coleoptera, Scarabaeoidea). *J Biogeogr* 30:1883–1897.

Schneck R, Micheels A, Mosbrugger V (2012) Climate impact of high northern vegetation: Late Miocene and present. *Int J Earth Sci (Geol Rundsch)* 101:323338. DOI10.1007/s00531-011-0652-4.

Scotese CR (2010) Point Tracker. PALEOMAP Project, University of Texas, Arlington; see www.scotese.com.

Smith SA (2009) Taking into account phylogenetic and divergence-time uncertainty in a parametric biogeographical analysis of the Northern Hemisphere plant clade Caprifolieae. *J Biogeogr* 36: 2324–2337.

Tiffney BH (1985a) Perspectives on the origin of the floristic similarity between eastern Asia and eastern North America. *J Arnold Arboretum* 66:73–94.

Tiffney BH (1985b) The Eocene North Atlantic land bridge: its importance in Tertiary and modern phytogeography of the Northern Hemisphere. *J Arnold Arboretum* 66:243–273.

Valencia-Barrera RM, Comtois P, Fernández-González D (2002) Bioclimatic indices as a tool in pollen forecasting. *Int J Biometeoro* l46:171–175.

Wang YJ, Susanna A, Von Raab-Straube E, Milne R, Liu, JQ (2009) Island-like radiation of Saussurea (Asteraceae: Cardueae) triggered by uplifts of the Qinghai–Tibetan Plateau. *Biol J Linn Soc* 97:893–903.

Wen J (1999) Evolution of eastern Asian and eastern North American disjunct distributions in flowering plants. *Annu Rev Ecol Evol Syst* 30:421–455.

Wen J, Ickert-Bond SM (2009) Evolution of the Madrean-Tethyan disjunctions and the North and South American amphitropical disjunctions in plants. *J Syst Evol* 47: 331–348.

Willis KJ, McElwain JC (2002) *The evolution of plants*. Oxford University Press, Oxford, UK.

Woillez M-N, Kageyama M, Krinner G, Noblet-Ducoudré N, Viovy N, Mancip M (2011) Impact of CO2 and climate on the Last Glacial Maximum vegetation: results from the ORCHIDEE/IPSL modelsClim. Past 7: 557–577.

Wolfe JA (1971) Tertiary climatic fluctuations and methods of analysis of Tertiary floras. *Palaeogeogr Palaeoclimatol Palaeoecol* 9:27–57.

Wolfe JA (1975) Some aspects of plant geography of the northern hemisphere during the Late cretaceous and Tertiary. *Ann Mo Bot Gard* 62: 264–279.

Xiang Q-Y, Soltis DE, Soltis PS, Manchester SR, Crawford DJ (2000) Timing the eastern Asian-eastern North American floristic disjunction: molecular clock corroborates paleontological estimates. *Mol Phylogenet Evol* 15: 462–472.

Zachos JC, Dickens GR, Zeebe RE (2008) An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature* 451: 279–283.

TABLES

Table S1. Effects of model uncertainty in biogeographic analysis: Changes in the estimation of range inheritance scenarios and parameter values when integrating additional sources of information: M1 (present distributions), M2 (present distributions, paleogeography), M3 (present and fossil distributions, paleogeography), M4 (present and fossil distributions, paleogeography, climatic niche). Node numbers correspond with the ancestral nodes of main clades depicted in Figure 2d. For each node, the range inheritance scenario ("Area split") with the highest relative likelihood value ("Rel. Prob") was reconstructed with the parametric likelihood model DEC (Ree & Smith 2008) over the maximum clade credibility tree from a MCMC BEAST analysis (Meseguer et al., 2013).

Node	M1			M2			M3			M4		
	Area split	Rel. Prob.	-lnL	Area split.	Rel. Prob.	-lnL	Area split	Rel. Prob.	-lnL	Area split	Rel. Prob.	-lnL
05	a|a	0.999	255.43	a|a	0.97	242.426	a|a	0.989	246.602	a|a	0.982	246.257
35	d|d	0.869	255.573	d|d	0.837	242.574	d|d	0.675	246.984	d|d	0.842	246.41
							b_d|d	0.181	248.299			
36	d|d	0.83	255.619	d|d	0.720	242.724	b_d|d	0.366	247.596	d|d	0.388	247.186
							d|d	0.175	248.331	b_d|d	0.327	247.355
							b|b	0.139	248.561	b|d	0.133	248.251
							b|d	0.135	248.588			
45	f|a	0.969	255.464	f|a	0.683	242.776	f|a	0.88	246.718	f|a	0.841	246.412
				a|a	0.199	244.01						
71	f|a_f	0.487	256.151	f|a_b_f	0.364	243.405	f|a_b_f	0.500	247.283	f|a_b_f	0.464	247.007
	f|a_b_f	0.275	256.722	f|a_f	0.351	243.443	f|a_f	0.277	247.872	f|a_f	0.306	247.423
	f|f	0.123	257.521									
137	a|a	0.923	255.513	a|a	0.976	242.42	a|a	0.988	246.603	a|a	0.989	246.25
139	a_f|a	0.411	256.321	a|a	0.369	243.392	a|a	0.369	247.586	a|a	0.518	246.896

	a|a	0.209	256.997	a_f|a	0.256	243.757	a_f|a	0.196	248.219	a|a_b a_f|a	0.133 0.112	248.252 248.428
141	a|a_d_f a|a_d	0.442 0.174	256.249 257.18	a|a_d_f a|a_d	0.272 0.256	243.697 243.758	a|a_b_ d a|a_b_f a|a_b a|b_f	0.388 0.277 0.147 0.119	247.598 247.933 248.563 248.773	a|a_b_d a|b_d	0.875 0.101	246.551 248.705

Table S2. Effects of phylogenetic uncertainty: Changes in the inference of ancestral ranges for the integrative M4 model when analysed with DEC over 1000 dated phylogenies sampled from the posterior distribution of the MCMC BEAST analysis in Meseguer et al. (2013). Node numbers correspond with those in Fig. 2d.

Clade/Node number	Ancestral Reconstruction	%
Clade B/36	D	0.74
	BD	0.26
D/71	ABF	0.67
	AF	0.29
	F	0.003
	BF	0.002
B/35	D	1
E/137	A	1
CDE/139	A	0.99
	AF	0.011
	ABF	0.002
C/45	A	0.37
	AF	0.63
A/5	A	1
Crown/141	ABD	1

Table S3. List of fossil occurrences. TS: Time Slice to which fossil remains have been attributed (1: Greenhouse, 2: Coldhouse).

TS	Period	subperiod	Fossil species	Lat.	Long.	RotLat	RotLong	Location	Reference
1	Eocene	Late	Hypericum antiquum Balueva et Nikit..	54.67	81.03	53.86	72.85	Uzhanikha, Novosibirskaya oblast, Russia	Arbuzova, 2005
1	Oligocene	Early	Hypericum sp.	41.48	1.33	37.13	-0.42	Catalonia, Spain	Cavagnetto & Anadón 1995
1	Oligocene	Early	Hypericum septestum Nikitin	56.48	84.97	55.93	76.29	Tomsk, Lagernyi Sad, Russia	Arbuzova, 2005
1	Oligocene	Late	Hypericum miocenicum Dorof. emend. Mai	51.05	15.00	46.97	11.31	Sachsen, Germany	Mai, 1997
1	Oligocene	Late	Hypericum septestum Nikitin	51.05	15.00	46.97	11.31	Sachsen, Germany	Mai, 1997
1	Oligocene	Late	Hypericum septestum Nikitin	51.07	14.10	46.98	10.48	Sachsen, Germany	Mai, 1997
1	Oligocene	Late	Hypericum septestum Nikitin	59.02	81.00	58.17	71.80	Tomsk, Russia	Dorofeev, 1963
1	Oligocene	Late	Hypericum septestum Nikitin	59.02	81.00	58.17	71.80	Tomsk, Russia	Dorofeev, 1963
1	Oligocene	Late	Hypericum septestum Nikitin	59.02	81.00	58.17	71.80	Tomsk, Russia	Dorofeev, 1963
1	Oligocene	Late	Hypericum septestum Nikitin	56.08	82.08	55.33	73.56	Tomsk, Russia	Dorofeev, 1963
1	Oligocene	Late	Hypericum septestum Nikitin	51.43	15.08	47.35	11.37	Gozdnica, Poland	Zastawniak et al., 1992
1	Oligocene		Hypericum bornense Mai	51.02	12.08	46.89	8.63	Sachsen, Germany	Mai & Walther, 1978
1	Oligocene		Hypericum septestum Nikitin	57.00	83.08	56.31	74.32	Tomsk, Russia	Dorofeev, 1963
1	Oligocene		Hypericum septestum Nikitin	57.00	83.08	56.31	74.32	Tomsk, Russia	Dorofeev, 1963
1	Miocene	Early	Hypericum cf. septestum Nikitin	52.63	83.00	51.97	75.18	Tomsk, Russia	Dorofeev, 1963
1	Miocene	Early	Hypericum cf. tertiaerum Nikitin	57.08	72.08	55.62	63.70	Omsk, Russia	Dorofeev, 1963
1	Miocene	Early	Hypericum holyi Friis	51.50	13.33	47.39	9.76	Lausizt, Germany	Mai, 2001
1	Miocene	Early	Hypericum septestum Nikitin	51.00	15.00	46.92	11.32	Zittau Basin, Czech Republic	Teodoris 2002
1	Miocene	Early	Hypericum septestum Nikitin	50.67	10.03	46.52	6.75	Hochrain bei Gerstengrund, Germany	Gumbel & Mai, 2002
1	Miocene	Early	Hypericum septestum Nikitin	56.08	74.08	54.77	65.83	Omsk, Russia	Dorofeev, 1963
1	Miocene	Early	Hypericum sp.	56.00	83.00	55.32	74.48	Tomsk, Russia	Dorofeev, 1963
1	Miocene	Early-Middle	Hypericum sp.	24.17	97.78	15.75	91.15	Longchuan County, Yunnan, China	Zhao et al., 2004
1	Miocene	Early-Middle	Hypericum septestum Nikitin	50.85	14.73	46.77	11.07	SE Saxony, SW Poland and N Bohemia	Kvacek & Teodiris, 2003
1	Miocene	Middle	Hypericum aff. ponticum Lipsky	43.55	23.07	39.66	19.18	Slavotin, Bulgaria	Palamarev et al., 2005
1	Miocene	Middle	Hypericum cf. androsaemum L.	47.08	28.30	43.35	23.88	Bursuc, Moldavia	Herpy, 1972

1	Miocene	Middle	Hypericum cf. balearicum L.	47.08	28.30	43.35	23.88	Bursuc, Moldavia	Herpy, 1972
1	Miocene	Middle	Hypericum cf. scabrum L.	47.08	28.30	43.35	23.88	Bursuc, Moldavia	Herpy, 1972
1	Miocene	Middle	Hypericum cf. septestum Nikitin	57.08	84.00	56.46	75.20	Tomsk, Russia	Dorofeev, 1963
1	Miocene	Middle	Hypericum holyi Friis	54.33	10.13	50.18	6.73	Jütland, Denmark	Mai, 2001, Friis, 1985
1	Miocene	Middle	Hypericum holyi Friis	51.10	13.15	46.99	9.61	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum miocenicum Dorof. emend. Mai	51.10	14.00	47.00	10.39	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum miocenicum Dorof. emend. Mai	51.08	13.08	46.97	9.54	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum miocenicum Dorof. emend. Mai	51.10	13.13	46.99	9.59	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum miocenicum Dorof. emend. Mai	51.10	13.15	46.99	9.61	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum miocenicum Dorof. emend. Mai	51.00	13.13	46.89	9.59	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum miocenicum Dorof. emend. Mai	51.07	14.10	46.98	10.48	Sachsen, Germany	Mai, 2001
1	Miocene	Middle	Hypericum miocenicum Dorof. emend. Mai	51.08	13.10	46.97	9.56	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum miocenicum Dorof. emend. Mai	51.10	14.07	47.01	10.46	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum septestum Nikitin	51.07	14.03	46.98	10.42	Sachsen, Germany	Mai, 2001
1	Miocene	Middle	Hypericum septestum Nikitin	51.07	14.03	46.98	10.42	Sachsen, Germany	Mai, 2001
1	Miocene	Middle	Hypericum septestum Nikitin	51.10	13.13	46.99	9.59	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum septestum Nikitin	51.10	13.15	46.99	9.61	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum septestum Nikitin	51.10	13.15	46.99	9.61	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum septestum Nikitin	51.10	13.15	46.99	9.61	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum septestum Nikitin	51.10	14.03	47.01	10.42	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum septestum Nikitin	51.10	13.15	46.99	9.61	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum septestum Nikitin	51.10	13.15	46.99	9.61	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum septestum Nikitin	51.05	14.10	46.96	10.48	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum septestum Nikitin	51.07	14.10	46.98	10.48	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum septestum Nikitin	51.07	14.13	46.98	10.51	Sachsen, Germany	Mai, 2001
1	Miocene	Middle	Hypericum septestum Nikitin	51.08	14.07	46.99	10.46	Sachsen, Germany	Mai, 2001
1	Miocene	Middle	Hypericum septestum Nikitin	43.70	22.98	39.81	19.09	Drenovets, Ga bare, Rouzhintsi, Bulgaria	Palamarev et al., 2005
1	Miocene	Middle	Hypericum sp.	40.38	-3.70	36.14	-5.19	Madrid, Spain	Barron et al., 2010
1	Miocene	Middle	Hypericum tertiaerum Nikitin	51.07	14.03	46.98	10.42	Sachsen, Germany	Mai, 2001
1	Miocene	Middle	Hypericum tertiaerum Nikitin	51.12	13.13	47.01	9.59	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum tertiaerum Nikitin	51.10	13.13	46.99	9.59	Brandenburg, Germany	Mai, 2001

1	Miocene	Middle	Hypericum tertiaerum Nikitin	51.10	13.13	46.99	9.59	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum tertiaerum Nikitin	51.10	13.15	46.99	9.61	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum tertiaerum Nikitin	51.10	13.15	46.99	9.61	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum tertiaerum Nikitin	51.10	13.15	46.99	9.61	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum tertiaerum Nikitin	51.10	13.15	46.99	9.61	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum tertiaerum Nikitin	51.10	13.15	46.99	9.61	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum tertiaerum Nikitin	51.10	13.15	46.99	9.61	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum tertiaerum Nikitin	51.10	13.15	46.99	9.61	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum tertiaerum Nikitin	51.05	14.10	46.96	10.48	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum tertiaerum Nikitin	51.07	14.10	46.98	10.48	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum tertiaerum Nikitin	51.10	14.07	47.01	10.46	Brandenburg, Germany	Mai, 2001
1	Miocene	Middle	Hypericum welzowense Mai	51.57	14.17	47.48	10.53	Welzow, Germany	Mai, 2001
2	Miocene	Late	Hypericum sp.	50.32	4.78	50.57	3.78	Entre-Sambre-et-Meuse, Belgium	Fairon-Demaret, 1996
2	Miocene	Late	Hypericum cf. acutum Moench	45.72	28.60	46.11	27.81	Tabaky, Ukraine	Herpy, 1972
2	Miocene	Late	Hypericum cf. holyi Friis	46.90	15.95	46.57	15.72	Fehring, East Styria, Austria	Meller & Hoffman, 2004
2	Miocene	Late	Hypericum cf. humifusum L.	44.83	9.95	44.56	9.71	Fidenza, Italy	Kovar-Eder et al., 2002
2	Miocene	Late	Hypericum cf. tertiaerum Nikitin	55.03	73.00	55.50	72.61	Omsk, Russia	Dorofeev., 1963
2	Miocene	Late	Hypericum cf. tertiaerum Nikitin	55.08	73.08	55.55	72.70	Omsk, Russia	Dorofeev., 1963
2	Miocene	Late	Hypericum cf. tertiaerum Nikitin	55.08	73.08	55.55	72.70	Omsk, Russia	Dorofeev., 1963
2	Miocene	Late	Hypericum cf. tertiaerum Nikitin	55.08	73.08	55.55	72.70	Omsk, Russia	Dorofeev., 1963
2	Miocene	Late	Hypericum cf. tertiaerum Nikitin	55.13	74.07	55.60	73.70	Omsk, Russia	Dorofeev., 1963
2	Miocene	Late	Hypericum cf. tertiaerum Nikitin	56.02	74.12	56.49	73.75	Omsk, Russia	Dorofeev., 1963
2	Miocene	Late	Hypericum cf. tertiaerum Nikitin	56.02	74.12	56.49	73.75	Omsk, Russia	Dorofeev., 1963
2	Miocene	Late	Hypericum cf. tertiaerum Nikitin	56.02	74.12	56.49	73.75	Omsk, Russia	Dorofeev., 1963
2	Miocene	Late	Hypericum cf. tertiaerum Nikitin	49.68	22.00	50.04	21.12	Sośnica, Poland	Łańcucka-S´rodoniowa, 1981
2	Miocene	Late	Hypericum holyi Friis	52.40	6.6	52.64	1.55	lower Rhenish Basin, Eschweiler, Germany	Van Der Burgh, 1987
2	Miocene	Late	Hypericum holyi Friis	51.53	13.88	51.84	12.92	Klettwitz, Germany	Mai, 2001
2	Miocene	Late	Hypericum kireevskiana	55.03	73.00	55.50	72.61	Omsk, Russia	Dorofeev., 1963
2	Miocene	Late	Hypericum miocenicum Dorof. emend. Mai	52.22	18.25	52.56	17.31	Konin, Gozdnica, Poland	Mai, 2001
2	Miocene	Late	Hypericum septestum Nikitin	57.00	75.02	57.46	74.67	Omsk, Russia	Dorofeev. 1963
2	Miocene	Late	Hypericum sp.	36.20	20.28	35.84	19.85	lower Rhenish Basin, Eschweiler, Germany	Van Der Burgh, 1987

2	Miocene	Late	Hypericum sp.	55.13	74.07	55.60	73.70	Omsk, Russia	Dorofeev., 1963
2	Miocene	Late	Hypericum sp. 1	44.83	9.95	44.56	9.71	close to Fidenza, Italy	Kovar-Eder et al., 2002
2	Miocene	Late	Hypericum sp. 2	44.83	9.95	44.56	9.71	close to Fidenza, Italy	Kovar-Eder et al., 2002
2	Miocene	Late	Hypericum tertiaerum Nikitin	51.43	15.08	51.75	14.13	Gozdnica, Poland	Dyjor et al. 1992
2	Miocene	Middle-Late	Hypericum septestum Nikitin	51.07	14.10	51.38	13.14	Brandenburg, Germany	Mai, 2001
2	Miocene	Middle-Late	Hypericum tertiaerum Nikitin	51.07	14.10	51.38	13.14	Brandenburg, Germany	Mai, 2001
2	Miocene		Hypericum sp.	45.72	28.60	46.11	27.81	Tabaky, Ukraine	Herpy, 1972
2	Miocene		Hypericum tambovicum P. Dorof.	44.75	40.40	45.19	39.71	Russia, Tambovski	Dorofeev. 1988
2	Miocene		Hypericum tanaiticum P. Dorof.	44.75	40.40	45.19	39.71	Russia, Tambovski	Dorofeev. 1988
2	Pliocene	Late	Hypericum xylosteifolium (Spach) N. Robson	43.00	40.98	43.44	40.30	Suhumi, Georgia	Arbuzova, 2005
2	Pliocene	Late	Hypericum androsaemun L.	50.83	10.02	51.12	9.04	Oberzella, Germany	Gumbel & Mai, 2004
2	Pliocene	Late	Hypericum foveolatum Dorofeev	50.60	10.15	50.89	9.18	Kaltensundheim, Germany	Gumbel & Mai, 2004
2	Pliocene	Late	Hypericum foveolatum Dorofeev	50.80	10.30	51.09	9.32	Barchfeld, Germany	Gumbel & Mai, 2004
2	Pliocene	Late	Hypericum tertiaerum Nikitin	50.60	10.15	50.89	9.18	Kaltensundheim, Germany	Gumbel & Mai, 2004
2	Pliocene	Late	Hypericum tertiaerum Nikitin	50.83	10.02	51.12	9.04	Oberzella, Germany	Gumbel & Mai, 2004
2	Pliocene	Middle	Hypericum cf. septestum Nikitin	44.80	7.82	44.55	7.59	Piedmont, Italy	Ciangherotti et al., 2007
2	Pliocene	Middle	Hypericum sp. 1	44.80	7.82	44.55	7.59	Piedmont, Italy	Ciangherotti et al., 2007
2	Pliocene	Pliocene	Hypericum sp.	5.17	74.50	1.74	72.65	Subachoque Valley, Cordillera Oriental, Colombia	Wijninga & Kurhy, 1990
2	Pliocene	Teglian, Reuverian, Cromerian	Hypericum cf. ascyron Linn.	50.38	8.07	50.65	7.09	Germany	Reid & Reid 1915
2	Pliocene		Hypericum sp.	11.00	40.50	10.49	39.78	Ethiopia	Bonnefille et al., 1987
2	Pliocene		Hypericum androsaemum L.	50.13	10.15	50.42	9.18	Thüringen, Germany	Mai & Walther, 1988
2	Pliocene		Hypericum calycinoides n sp.	51.08	11.05	51.37	10.07	Thüringen, Germany	Mai & Walther, 1988
2	Pliocene		Hypericum cf. danicum Friis	44.15	8.00	43.90	7.75	Piemonte, Italy	Mai, 1995
2	Pliocene		Hypericum cf. perforatum L.	50.10	10.03	50.39	9.06	Thüringen, Germany	Mai & Walther, 1988
2	Pliocene		Hypericum foveolatum Dorofeev	52.32	30.93	52.72	30.09	Gómel, Belarus	Velichkevich & Zastawniak,

									2003
2	Pliocene		Hypericum hirsutum L.	50.15	11.03	50.44	10.07	Thüringen, Germany	Mai & Walther, 1988
2	Pliocene		Hypericum miocenicum Dorof. emend. Mai	50.98	10.83	51.27	9.85	Thüringen, Germany	Mai & Walther, 1988
2	Pliocene		Hypericum perforatum L.	50.15	11.03	50.44	10.07	Thüringen, Germany	Mai & Walther, 1988
2	Pliocene		Hypericum perforatum L.	51.63	39.18	52.07	38.43	Russia, Voronej	Arbuzova, 2005
2	Pliocene		Hypericum septestum Nikitin	45.08	7.72	44.83	7.49	Piedmont, Italy	Martinetto et al., 2007
2	Pliocene		Hypericum septestum Nikitin	50.07	12.37	50.37	11.42	Bohemian Massif, Czech republic	Kvacek & Teodiris, 2003
2	Pliocene		Hypericum sp.	45.08	7.72	44.83	7.49	Piedmont, Italy	Martinetto et al., 2007
2	Pliocene		Hypericum sp A	45.08	7.72	44.83	7.49	Piedmont, Italy	Martinetto et al., 2007
2	Pliocene		Hypericum sp.	52.32	30.93	52.72	30.09	Gómel, Belarus	Velichkevich & Zastawniak, 2003
2	Pliocene		Hypericum tertiaerum Nikitin	51.63	39.18	52.07	38.43	Voronej, Russia	Nikitin, 1957
2	Pliocene		Hypericum tertiaerum Nikitin	49.45	20.28	49.08	20.07	Mizerna in den Vorkarpaten, Poland	Mai, 2001
2	Pliocene		Hypericum tertiaerum Nikitin	51.33	6.13	51.59	5.12	Tegelen, Holland	Nikitin, 1957
2	Pliocene		Hypericum tertiaerum Nikitin	51.28	6.08	51.54	5.07	Reuver, Holland	Nikitin, 1957
2	Pliocene		Hypericum tertiaerum Nikitin	52.32	30.93	52.72	30.09	Gómel, Belarus	Velichkevich & Zastawniak, 2003
2	Pliocene		Hypericum tertiaerum Nikitin	50.98	10.83	51.27	9.85	Thüringen, Germany	Mai & Walther, 1988
2	Pliocene		Hypericum tetrapterum Fries	50.15	11.03	50.44	10.07	Thüringen, Germany	Mai & Walther, 1988
2	Pliocene		Hypericum tetrapterum Fries	50.13	10.15	50.42	9.18	Thüringen, Germany	Mai & Walther, 1988
2	Pliocene		Hypericum cf. tertiaerum Nikitin	51.40	15.17	51.72	14.22	Ruszów, Poland	Baranowska-Zarzycka, 1988
2	Pliocene		Hypericum foveolatum Dorofeev	53.88	30.40	54.28	29.54	Dvorets on the Dnieper, Belarus	Dorofeev, 1986a
2	Pliocene, Quaternary		Hypericum cf. tertiaerum Nikitin	51.63	39.20	52.07	38.45	Voronezh, Krivobor'e, Russia	Nikitin, 1957
2	Pliocene- Pleistocene		Hypericum sp.	4.58	-74.07	4.18	-72.13	High plain of Bogotá, Colombia	Hammen, 1974
-	Pleistocene		Hypericum sp.	-3.00	30.00	-3.54	28.97	Burundi	Bonnefille & Riollet, 1987
-	Pleistocene		Hypericum sp.	-3.47	29.57	-4.01	28.55	Burundi	Bonnefille et al., 1992
-	Pleistocene		Hypericum sp.	37.35	140.63	39.09	138.68	Fukushima, Japan	Tsukada 1985
-	Pleistocene		Hypericum sp.	-3.00	35.05	-3.51	34.03	Tanzania	Hay, 1990
-	Pleistocene		Hypericum sp.	-3.00	35.05	-3.51	34.03	Tanzania	Tobias, 1991
-	Pleistocene		Hypericum virginicum (L.) Rafin.	42.50	78.50	42.96	78.11	Sardinia Village, New York	Miller & Calkin, 1990

SUPPLEMENTARY INFORMATION FIGURE CAPTIONS

Figure S1: Worldwide geographic projection of the climatic optimum of "extant" *Hypericum* species (based on present distributions) over six climate simulations representing major warming or cooling events in Earth history, using the inverse normalized Mahalanobis Distance (MD) as a measure of climatic favourability: a) Early Eocene (with CO2 level 1120 ppm, b) Late Eocene (560 ppm), c) Miocene 400 ppm, d) Miocene 280 ppm, e) Pliocene 560 ppm, and f) Preindustrial 280 ppm.

Figure S2: Worldwide geographic projection of the climatic optimum of "ancestral" *Hypericum* species (based on fossil distributions) over six climate simulations representing major warming or cooling events in Earth history, using the inverse normalized Mahalanobis Distance (MD) as a measure of climatic favourability: a, g) Early Eocene, b, h) Late Eocene, c, i) Miocene, d, j) Late Miocene, e, k) Pliocene, and f, l) Preindustrial world. "Greenhouse" fossil lineages refer to fossil records between Late Eocene and Early-Mid Miocene (boundary at the Late Miocene cooling event, ca. 11 Ma). "Coldhouse" fossil lineages refer to fossil records between Late Miocene and Pliocene (Pleistocene fossil records were excluded from the analyses).

Figure S3: Bar-and-whisker plots showing median values and 25% and 75% quartiles of the Mahalanobis Distance for all Earth pixels with present (A) or fossil (B) occurrences, according to the optimum values estimated from present observations and greenhouse or coldhouse fossil data. Values with the same letter do not differ significantly.

Figure S4: Inter-temporal transferability: A) Greenhouse fossil optimum projected onto a "coldhouse" climate layer (Pliocene 560 ppm). B) Coldhouse fossil optimum projected onto a "greenhouse" climate layer (Eocene 560 ppm). C-D) Similar analyses repeated excluding the intermediate Miocene fossils (see text), and with dots representing fossil locations (red dots: greenhouse; blue dots: coldhouse fossils).

Figure S5: Variability of the ENFA (Ecological Niche Factor Analysis) factors based on current observations. a) First ENFA factor ("tropicality"). b) Second ENFA factor ("aridity"). Factor score values of climatic variables are given in SI Text.

Figure S6: Reconstruction of the biogeographic history of *Hypericum* based on extant species distributions (only present occurrences; M1 model, see text). Ancestral ranges and biogeographic events were reconstructed with the parametric likelihood model DEC (Ree & Smith 2008) over the maximum clade credibility tree from a MCMC BEAST analysis. Nodal pie charts represent relative probabilities for alternative ancestral distributions, whereas square splits indicate range inheritance scenarios showing how the ancestral range with the highest likelihood was divided at the speciation event; colours correspond with the discrete areas in the inset map. Arrows along phylogenetic branches represent single dispersal events, while red crosses show extinction events inferred by the model. Coloured circles before the species name give present ranges.

Figure S7: Reconstruction of the biogeographic history of *Hypericum* (M2 model, see text) based on extant species distributions but adding a palaeogeographic model which stratifies the phylogeny into time slices (TS) and assigns different probabilities of dispersal between areas according to the geographic configuration of continents (and general climatic and vegetation belts) during Cenozoic (Tertiary) history. See text for more details on scaling of dispersal rates. Shadow grey vertical bars indicate the limit between times slices (TS). All other conventions like in Fig. S6.

Figure S8: Reconstruction of the biogeographic history of *Hypericum* based on extant-species distributions and palaeogeographic information (as in SI Fig. 7), but also adding information from fossil ranges (M3, see text). In particular, the ancestral range of crown node *Hypericum* (node number 141) was constrained to include the range of the oldest *Hypericum* fossil, *H. antiquum* from Siberia, i.e., only the Eastern Palearctic or widespread ancestral ranges including this area were accepted as possible ancestral states in the biogeographic model. All other conventions like in Fig. S6.

Figure S9: Reconstruction of the biogeographic history of *Hypericum* based on extant species distributions and fossil ranges (like in Fig. S8), but also including information on the favourability of the world to past *Hypericum* climatic requirements as inferred from fossil-based ENM models. This was achieved by a) constraining the crown group ancestral range to include the Eastern Palearctic and the Nearctic regions, the latter is not represented in the fossil record but was predicted by ENM models to be part of the ancestral distribution. b) Modifying the dispersal rates in the paleogeographic model to reflect not only "geographic connectivity" (e.g., the presence of a land bridge between two continents) but also "ecological connectivity": ancestral climatic preferences and predicted dispersal routes as inferred from fossil-based ENM models (SI Figure 4; see text for further explanation). To simplify the interpretations, the continuous MD representation in Fig. S3 has been transformed into four colour categories: blue (pixels with MD < 5), green (MD = 5-10), yellow (MD = 10–15), and red (MD > 15). Empty circles represent selected fossil localities. Global mean temperature curve obtained from Zachos et al. (2001). All other conventions like in Fig. S6.

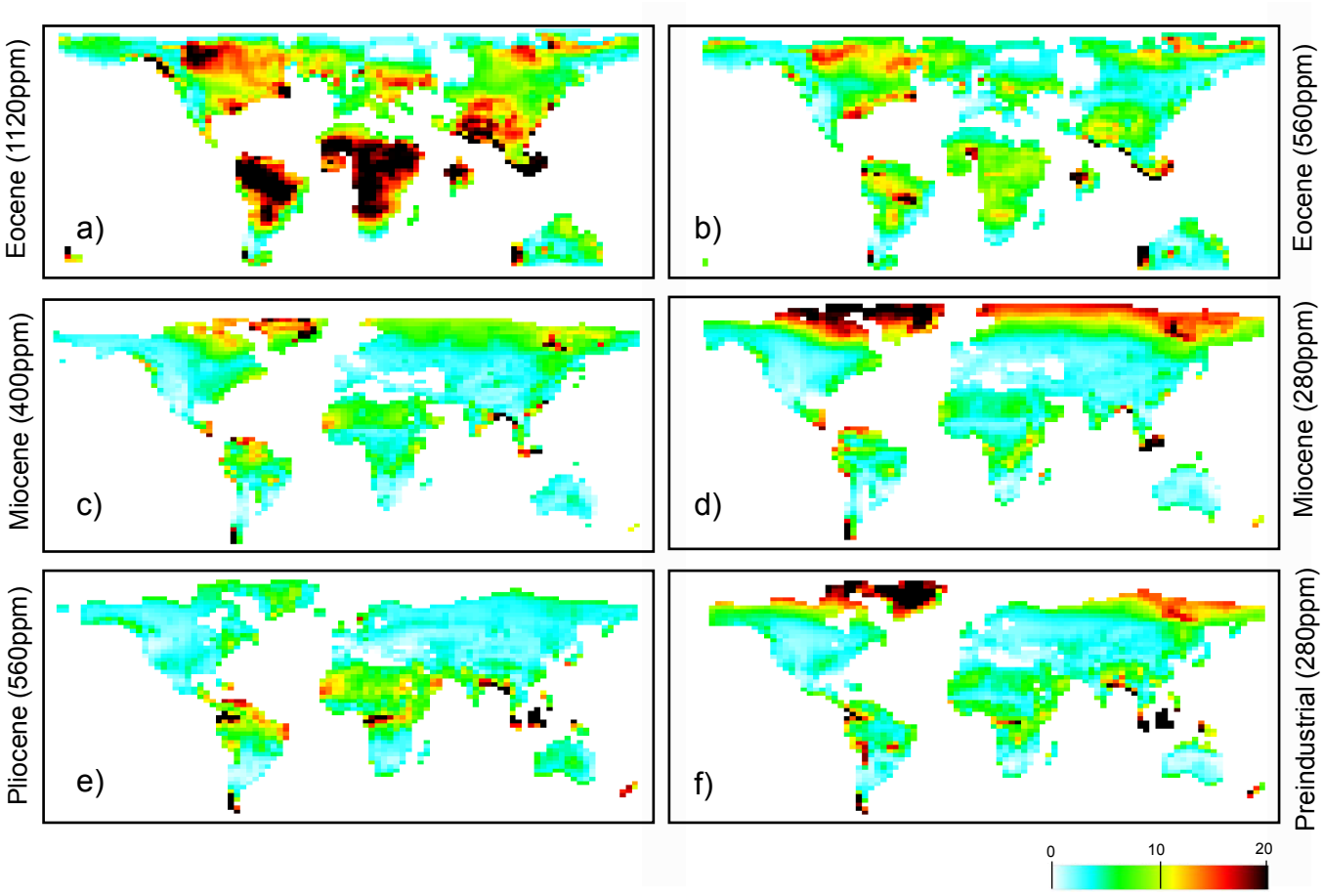

Eocene (1120ppm)
a)
b)
Eocene (560ppm)
Miocene (400ppm)
c)
d)
Miocene (280ppm)
Pliocene (560ppm)
e)
f)
Preindustrial (280ppm)
0
10
20

Fig S1

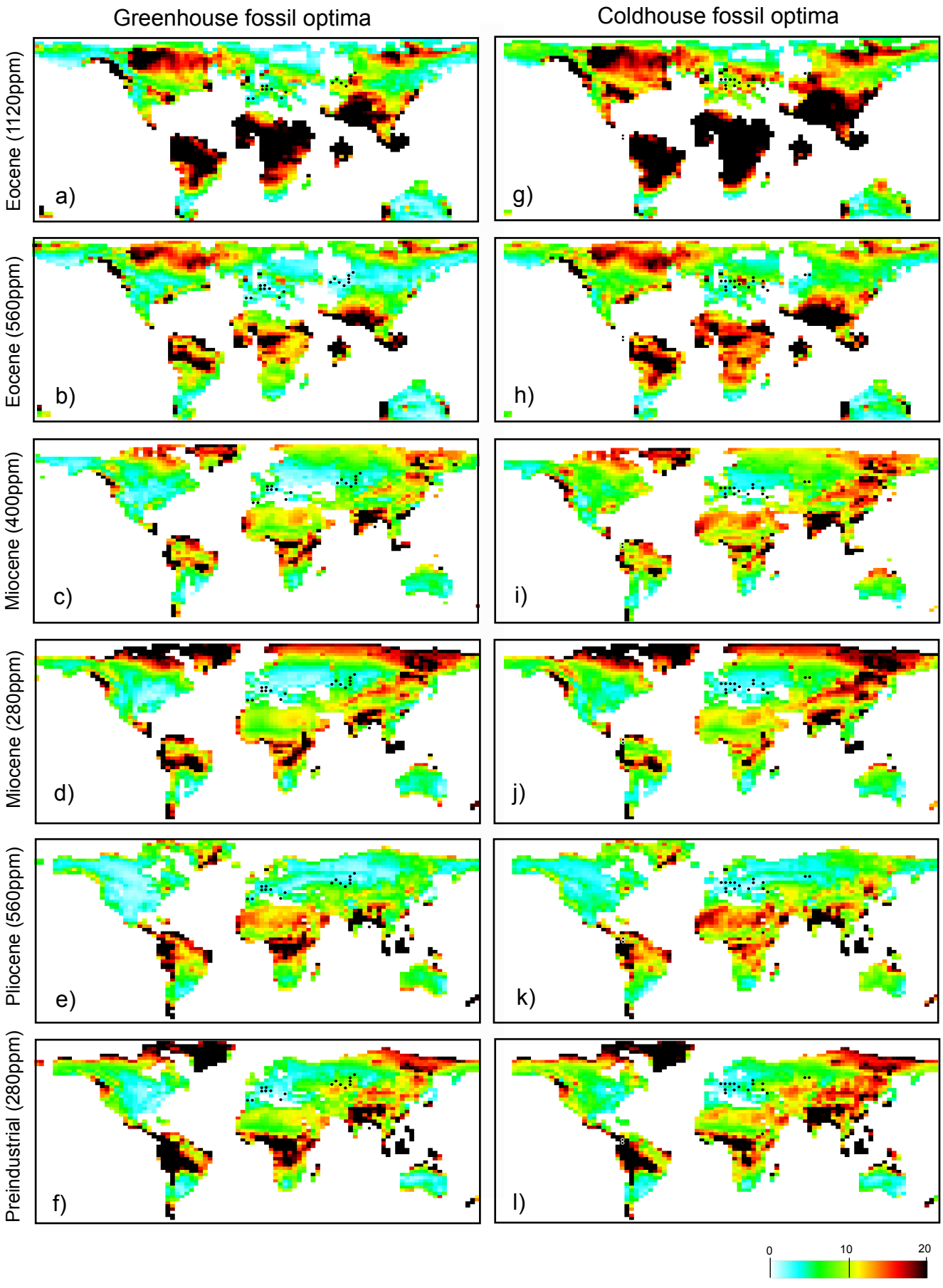

Greenhouse fossil optima
Coldhouse fossil optima
Eocene (1120ppm)
Eocene (560ppm)
Miocene (400ppm)
Miocene (280ppm)
Pliocene (560ppm)
Preindustrial (280ppm)
a)
b)
c)
d)
e)
f)
g)
h)
i)
j)
k)
l)
0
10
20

Fig S2

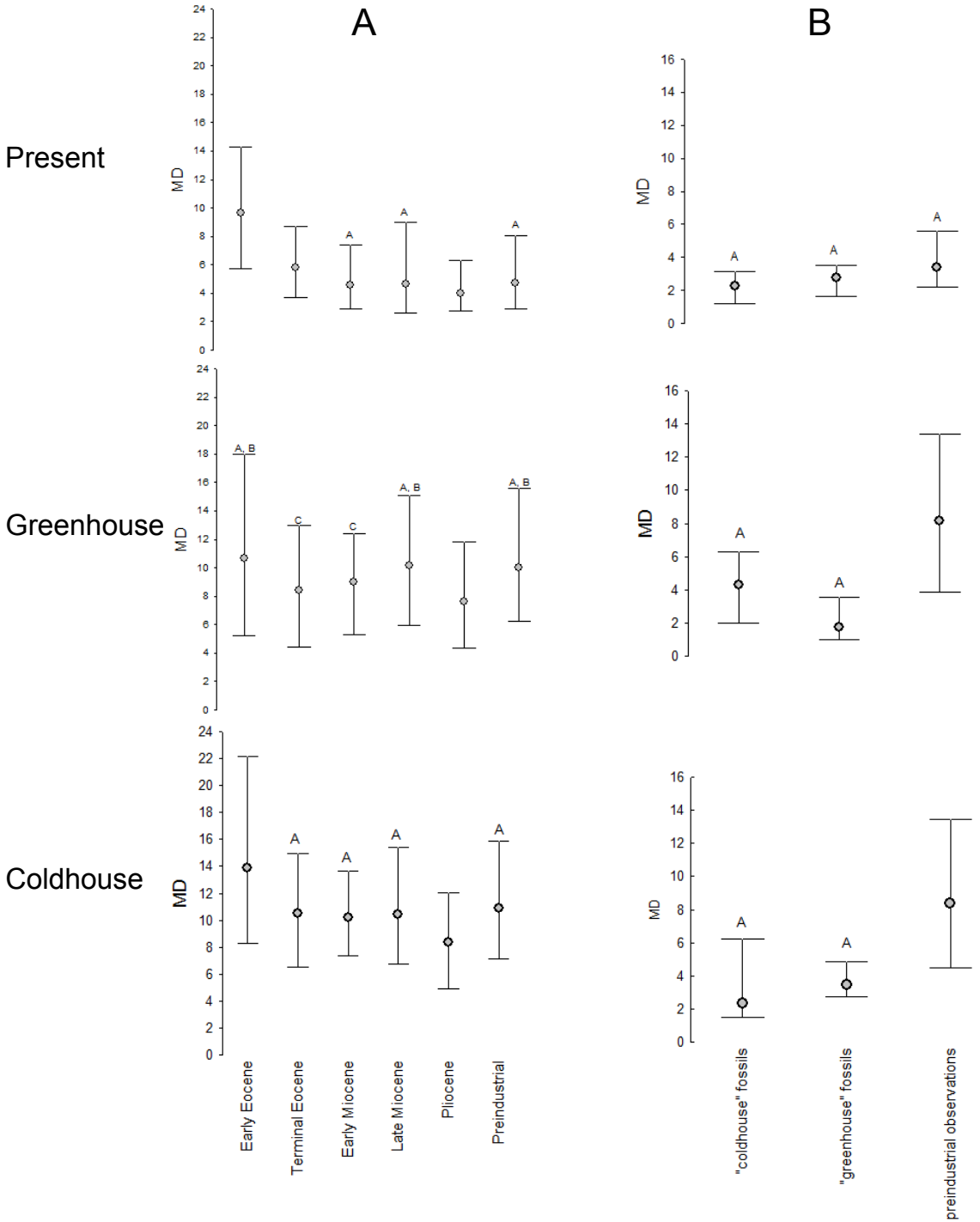

A
B
Present
Greenhouse
Coldhouse
MD
A
A
A
A
A
A
A, B
C
C
A, B
A, B
A
A
A
A
A
A
A
A
Early Eocene
Terminal Eocene
Early Miocene
Late Miocene
Pliocene
Preindustrial
"coldhouse" fossils
"greenhouse" fossils
preindustrial observations

Fig S3

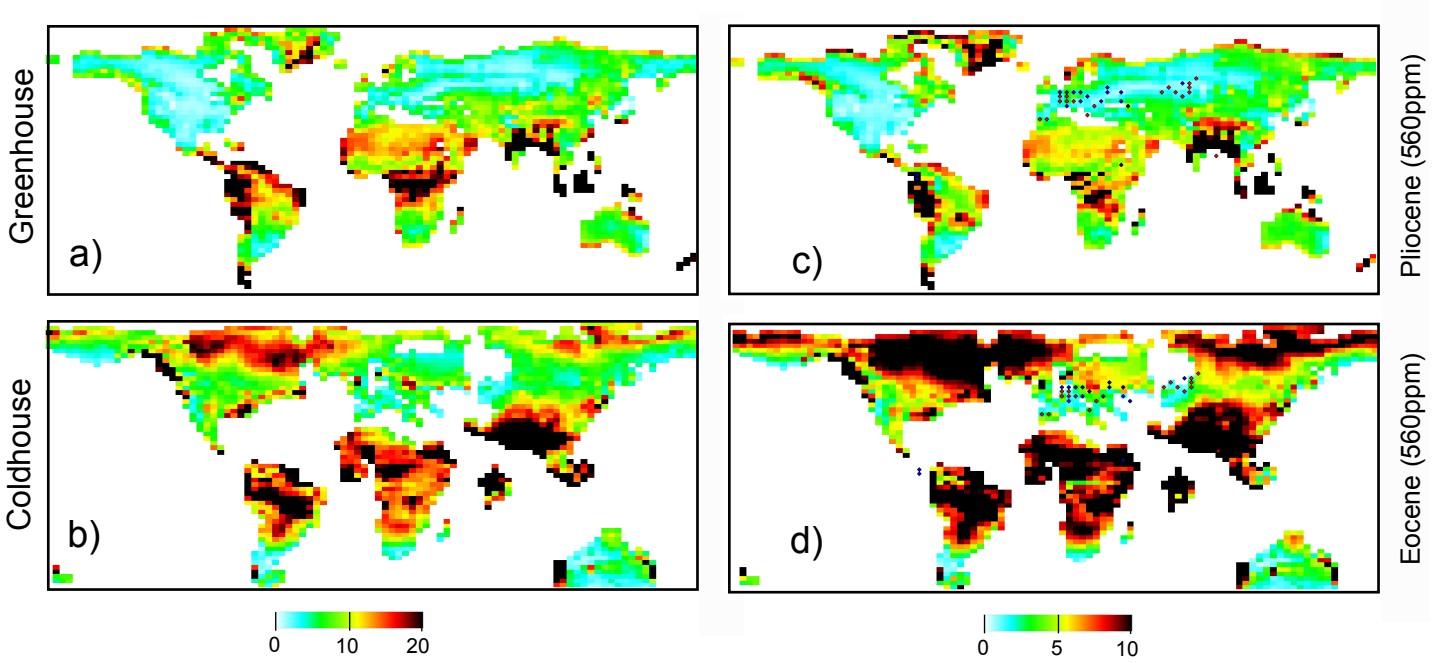
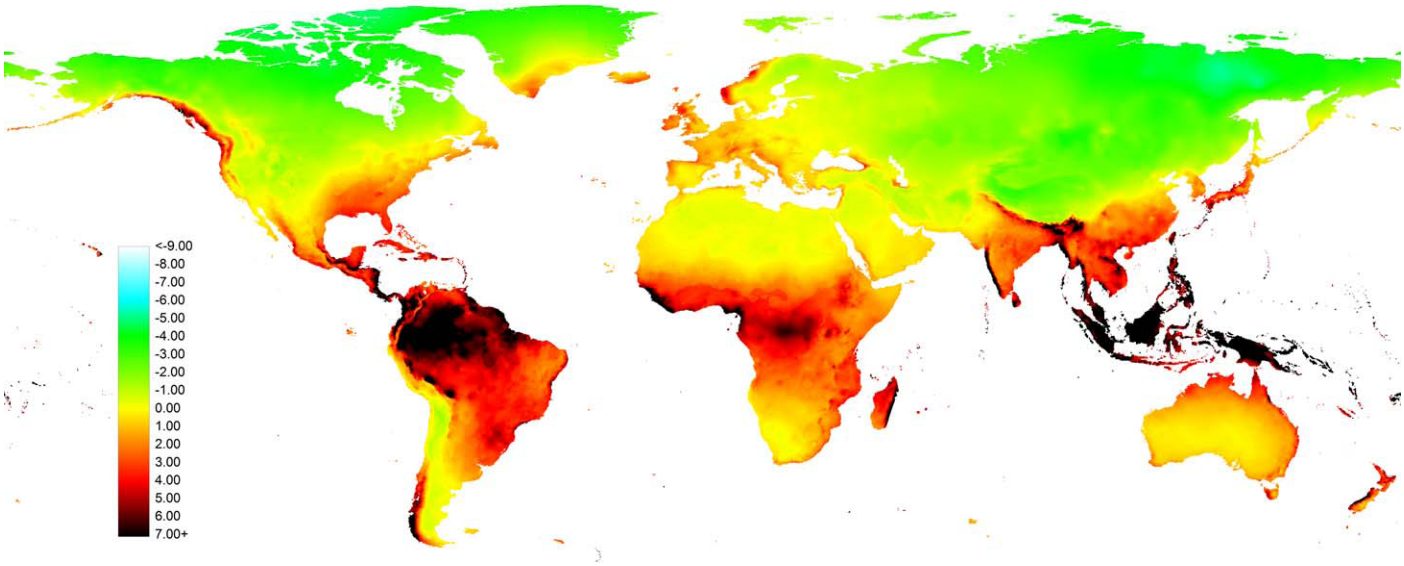

Greenhouse
a)
c)
Pliocene (560ppm)
Coldhouse
b)
d)
Eocene (560ppm)
0
10
20
0
5
10

Fig S4

Factor 1



Factor 2

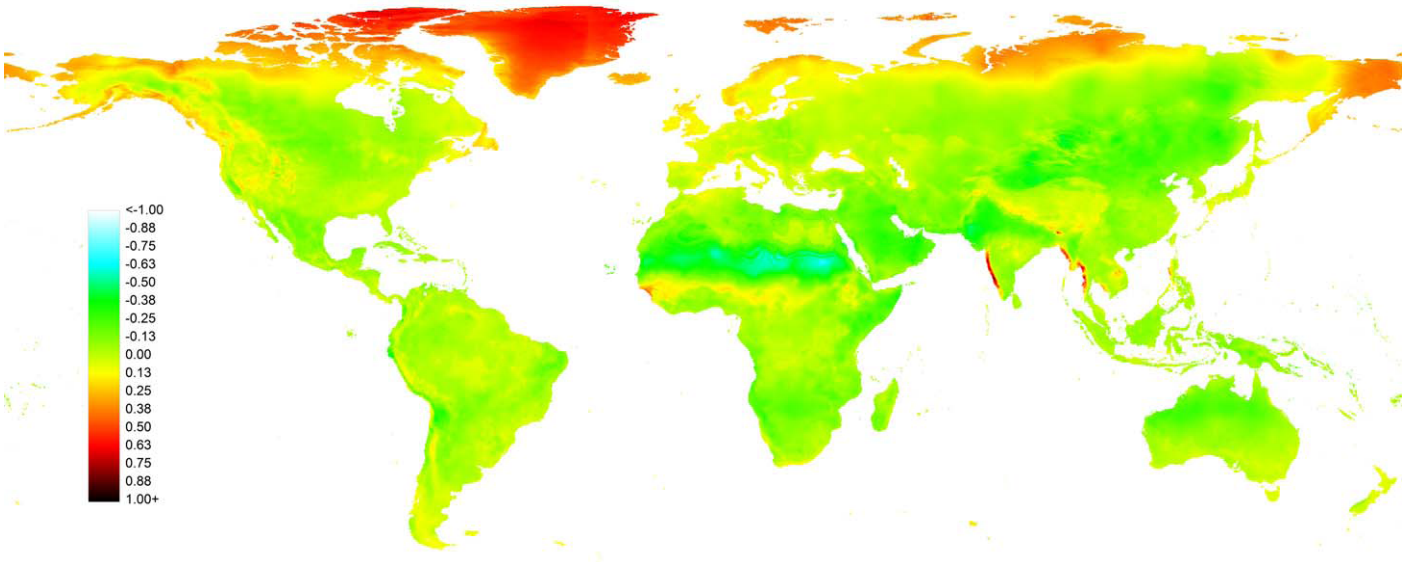


Fig S5

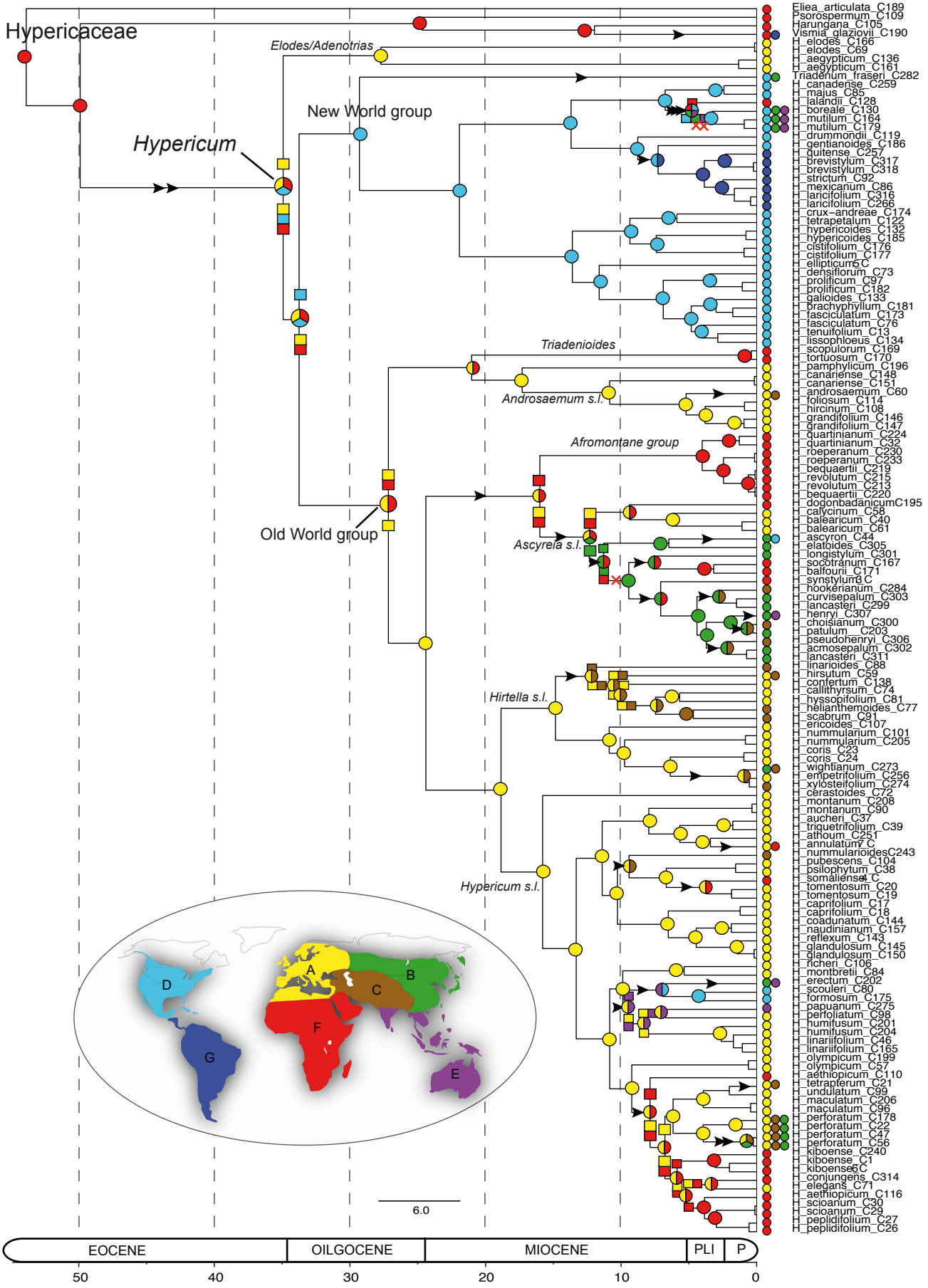

Hypericaceae
Elodes/Adenotrias
New World group
Hypericum
Triadenioides
Androsaemum s.l.
Afromontane group
Old World group
Ascyreia s.l.
Hirtella s.l.
Hypericum s.l.
A
B
C
D
E
F
G
6.0
EOCENE
OILGOCENE
MIOCENE
PLI
P
50
40
30
20
10
0

Fig S6

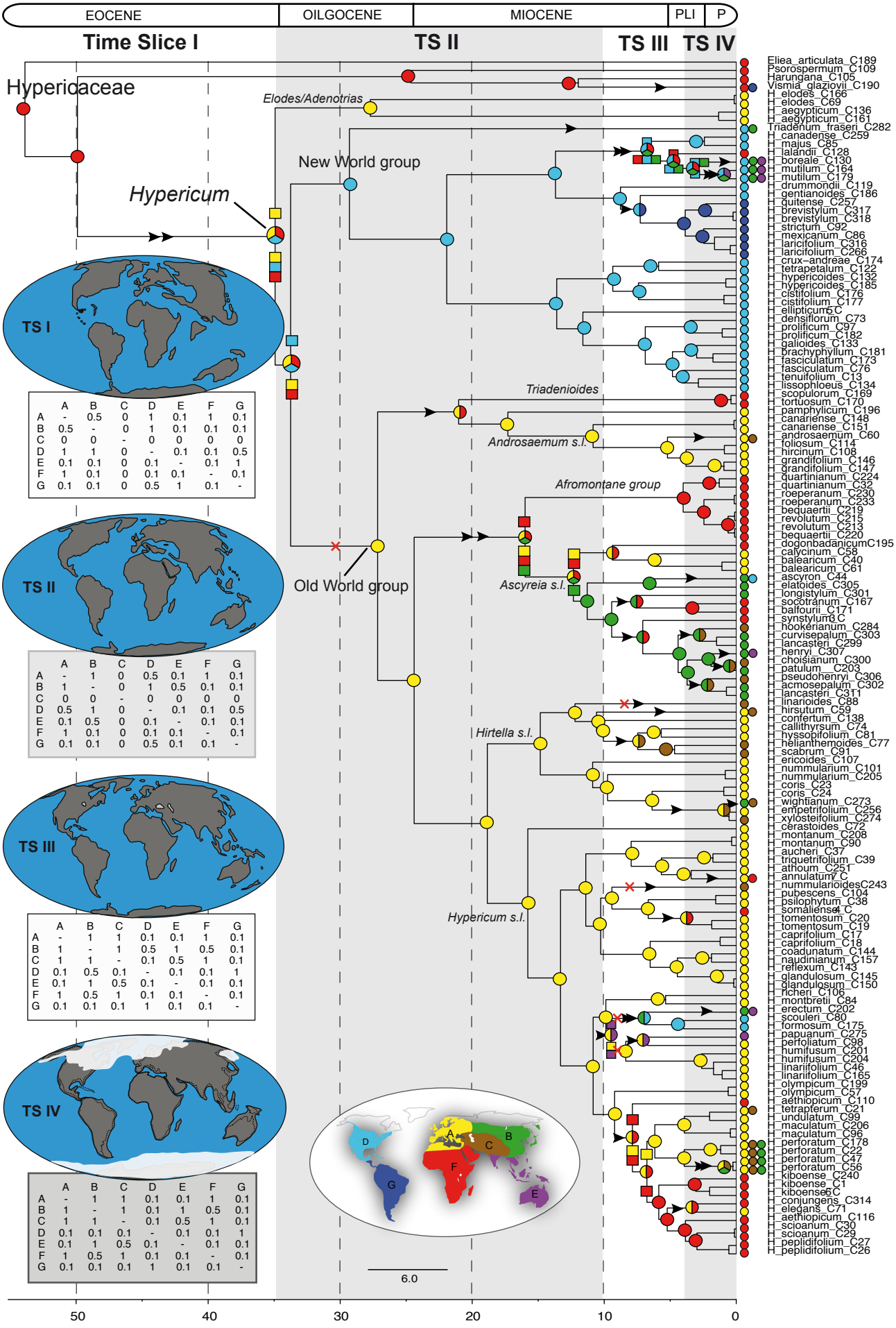

EOCENE
OILGOCENE
MIOCENE
PLI
P
Time Slice I
TS II
TS III
TS IV
Hypericaceae
Elodes/Adenotrias
New World group
Hypericum
Triadenioides
Androsaemum s.l.
Afromontane group
Ascyreia s.l.
Old World group
Hirtella s.l.
Hypericum s.l.
TS I
TS II
TS III
TS IV
A B C D E F G
A - 0.5 0 1 0.1 1 0.1
B 0.5 - 0 1 0.1 0.1 0.1
C 0 0 - 0 0 0 0
D 1 1 0 - 0.1 0.1 0.5
E 0.1 0.1 0 0.1 - 0.1 1
F 1 0.1 0 0.1 0.1 - 0.1
G 0.1 0.1 0 0.5 1 0.1 -
A B C D E F G
A - 1 0 0.5 0.1 1 0.1
B 1 - 0 1 0.5 0.1 0.1
C 0 0 - 0 0 0 0
D 0.5 1 0 - 0.1 0.1 0.5
E 0.1 0.5 0 0.1 - 0.1 0.1
F 1 0.1 0 0.1 0.1 - 0.1
G 0.1 0.1 0 0.5 0.1 0.1 -
A B C D E F G
A - 1 1 0.1 0.1 1 0.1
B 1 - 1 0.5 1 0.5 0.1
C 1 1 - 0.1 0.5 1 0.1
D 0.1 0.5 0.1 - 0.1 0.1 1
E 0.1 1 0.5 0.1 - 0.1 0.1
F 1 0.5 1 0.1 0.1 - 0.1
G 0.1 0.1 0.1 1 0.1 0.1 -
A B C D E F G
A - 1 1 0.1 0.1 1 0.1
B 1 - 1 0.1 1 0.5 0.1
C 1 1 - 0.1 0.5 1 0.1
D 0.1 0.1 0.1 - 0.1 0.1 1
E 0.1 1 0.5 0.1 - 0.1 0.1
F 1 0.5 1 0.1 0.1 - 0.1
G 0.1 0.1 0.1 1 0.1 0.1 -
A
B
C
D
E
F
G
6.0
50
40
30
20
10
0
Eliea_articulata_C189
Psorospermum_C109
Harungana_C105
Vismia_glaziovii_C190
H_elodes_C166
H_elodes_C69
H_aegypticum_C136
H_aegypticum_C161
Triadenum_fraseri_C282
H_canadense_C259
H_majus_C85
H_lalandii_C128
H_boreale_C130
H_mutilum_C164
H_mutilum_C179
H_drummondii_C119
H_gentianoides_C186
H_quitense_C257
H_brevistylum_C317
H_brevistylum_C318
H_strictum_C92
H_mexicanum_C86
H_laricifolium_C316
H_laricifolium_C266
H_crux-andreae_C174
H_tetrapetalum_C122
H_hypericoides_C132
H_hypericoides_C185
H_cistifolium_C176
H_cistifolium_C177
H_ellipticum5C
H_densiflorum_C73
H_prolificum_C97
H_prolificum_C182
H_galioides_C133
H_brachyphyllum_C181
H_fasciculatum_C173
H_fasciculatum_C76
H_tenuifolium_C13
H_lissophloeus_C134
H_scopulorum_C169
H_tortuosum_C170
H_pamphylicum_C196
H_canariense_C148
H_canariense_C151
H_androsaemum_C60
H_foliosum_C114
H_hircinum_C108
H_grandifolium_C146
H_grandifolium_C147
H_quartinianum_C224
H_quartinianum_C32
H_roeperanum_C230
H_roeperanum_C233
H_bequaertii_C219
H_revolutum_C215
H_revolutum_C213
H_bequaertii_C220
H_dogonbadanicumC195
H_calycinum_C58
H_balearicum_C40
H_balearicum_C61
H_ascyron_C44
H_elatoides_C305
H_longistylum_C301
H_socotranum_C167
H_balfourii_C171
H_synstylum3C
H_hookerianum_C284
H_curvisepalum_C303
H_lancasteri_C299
H_henryi_C307
H_choisianum_C300
H_patulum__C203
H_pseudohenryi_C306
H_acmosepalum_C302
H_lancasteri_C311
H_linarioides_C88
H_hirsutum_C59
H_confertum_C138
H_callithyrsum_C74
H_hyssopifolium_C81
H_helianthemoides_C77
H_scabrum_C91
H_ericoides_C107
H_nummularium_C101
H_nummularium_C205
H_coris_C23
H_coris_C24
H_wightianum_C273
H_empetrifolium_C256
H_xylosteifolium_C274
H_cerastoides_C72
H_montanum_C208
H_montanum_C90
H_aucheri_C37
H_triquetrifolium_C39
H_athoum_C251
H_annulatum7C
H_nummularioidesC243
H_pubescens_C104
H_psilophytum_C38
H_somalense4C
H_tomentosum_C20
H_tomentosum_C19
H_caprifolium_C17
H_caprifolium_C18
H_coadunatum_C144
H_naudinianum_C157
H_reflexum_C143
H_glandulosum_C145
H_glandulosum_C150
H_richeri_C106
H_montbretii_C84
H_erectum_C202
H_scouleri_C60
H_formosum_C175
H_papuanum_C275
H_perfoliatum_C98
H_humifusum_C201
H_humifusum_C204
H_linariifolium_C46
H_linariifolium_C165
H_olympicum_C199
H_olympicum_C57
H_aethiopicum_C110
H_tetrapterum_C21
H_undulatum_C99
H_maculatum_C206
H_maculatum_C96
H_perforatum_C178
H_perforatum_C22
H_perforatum_C47
H_perforatum_C56
H_kiboense_C240
H_kiboense_C1
H_kiboense6C
H_conjungens_C314
H_elegans_C71
H_aethiopicum_C116
H_scioanum_C30
H_scioanum_C29
H_peplidifolium_C27
H_peplidifolium_C26

Fig S7

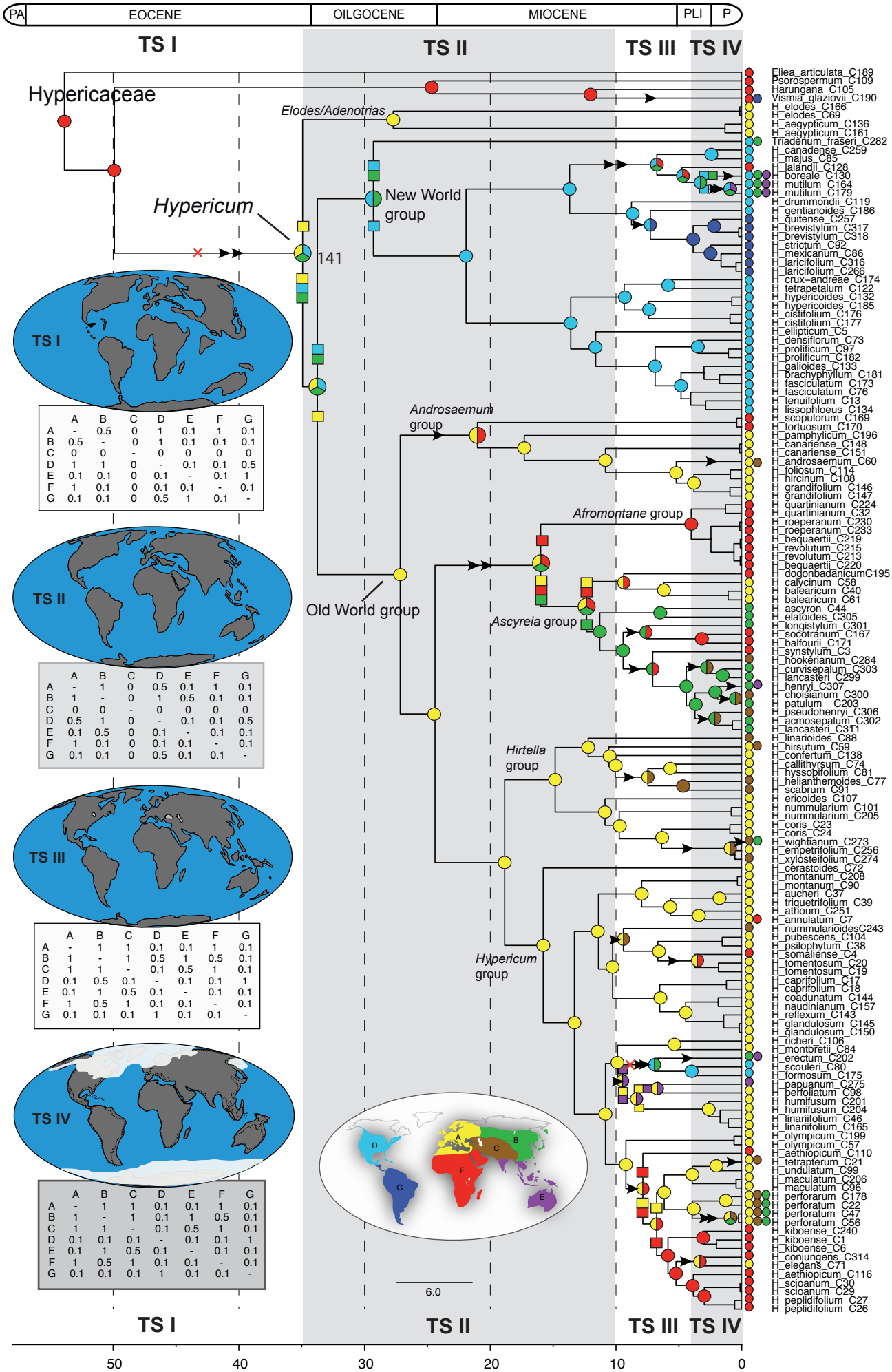

PA
EOCENE
OILGOCENE
MIOCENE
PLI
P
TS I
TS II
TS III
TS IV
Hypericaceae
Elodes/Adenotrias
Hypericum
141
New World group
Old World group
Androsaemum group
Afromontane group
Ascyreia group
Hirtella group
Hypericum group
TS I
TS II
TS III
TS IV
A B C D E F G
A - 0.5 0 1 0.1 1 0.1
B 0.5 - 0 1 0.1 0.1 0.1
C 0 0 - 0 0 0 0
D 1 1 0 - 0.1 0.1 0.5
E 0.1 0.1 0 0.1 - 0.1 1
F 1 0.1 0 0.1 0.1 - 0.1
G 0.1 0.1 0 0.5 1 0.1 -
A B C D E F G
A - 1 0 0.5 0.1 1 0.1
B 1 - 0 1 0.5 0.1 0.1
C 0 0 - 0 0 0 0
D 0.5 1 0 - 0.1 0.1 0.5
E 0.1 0.5 0 0.1 - 0.1 0.1
F 1 0.1 0 0.1 0.1 - 0.1
G 0.1 0.1 0 0.5 0.1 0.1 -
A B C D E F G
A - 1 1 0.1 0.1 1 0.1
B 1 - 1 0.5 1 0.5 0.1
C 1 1 - 0.1 0.5 1 0.1
D 0.1 0.5 0.1 - 0.1 0.1 1
E 0.1 1 0.5 0.1 - 0.1 0.1
F 1 0.5 1 0.1 0.1 - 0.1
G 0.1 0.1 0.1 1 0.1 0.1 -
A B C D E F G
A - 1 1 0.1 0.1 1 0.1
B 1 - 1 0.1 1 0.5 0.1
C 1 1 - 0.1 0.5 1 0.1
D 0.1 0.1 0.1 - 0.1 0.1 1
E 0.1 1 0.5 0.1 - 0.1 0.1
F 1 0.5 1 0.1 0.1 - 0.1
G 0.1 0.1 0.1 1 0.1 0.1 -
A
B
C
D
E
F
G
6.0
50
40
30
20
10
0

Fig S8

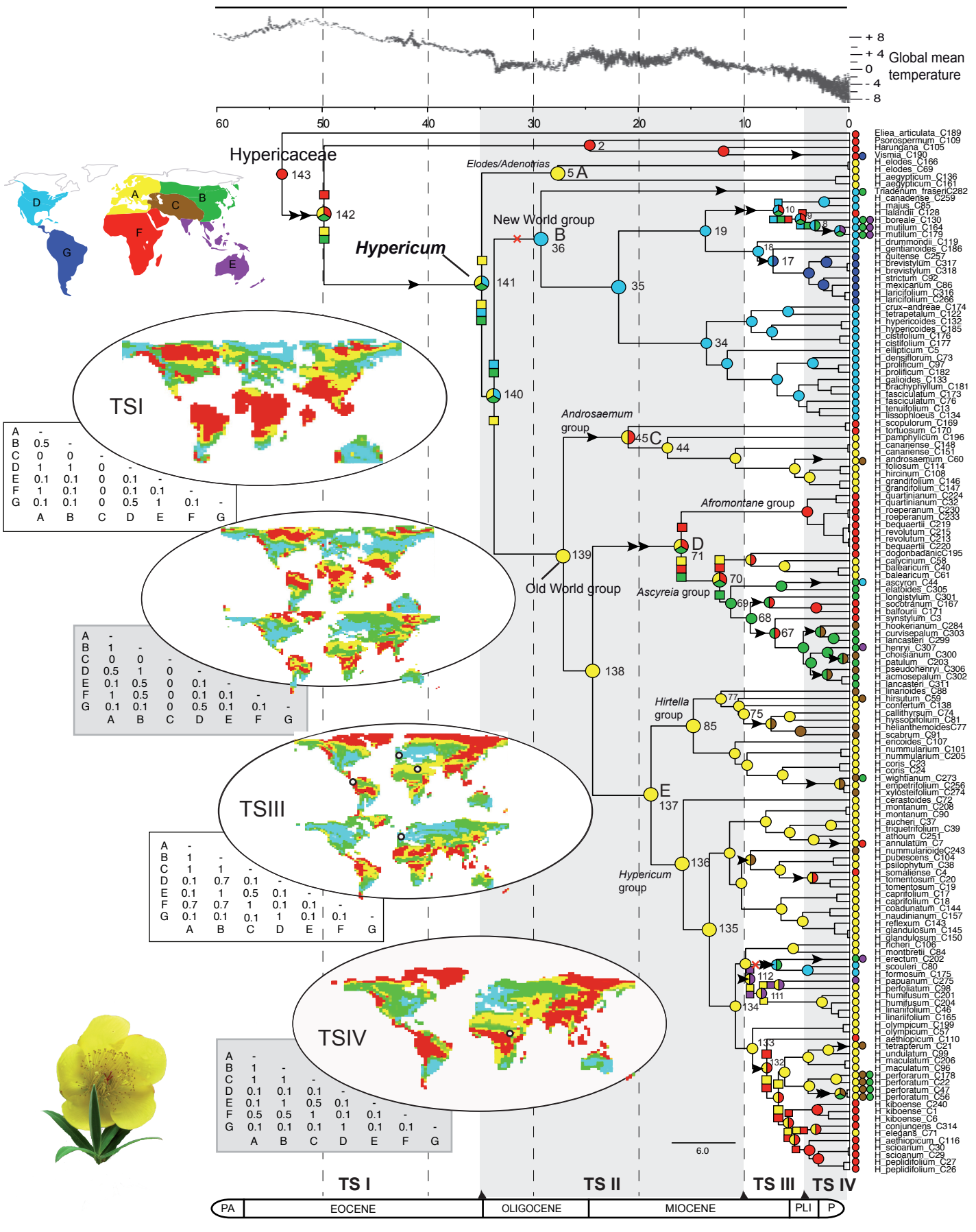

+ 8
+ 4
0
- 4
- 8
Global mean temperature
60
50
40
30
20
10
0
Hypericaceae
143
142
Hypericum
141
Elodes/Adenotrias
5 A
2
New World group
B
36
19
18
17
10
9
8
35
34
140
Androsaemum group
45 C
44
Afromontane group
D
71
70
69
68
67
139
Old World group
Ascyreia group
138
Hirtella group
85
77
75
E
137
Hypericum group
136
135
134
133
132
112
111
6.0
TSI
A -
B 0.5 -
C 0 0 -
D 1 1 0 -
E 0.1 0.1 0 0.1 -
F 1 0.1 0 0.1 0.1 -
G 0.1 0.1 0 0.5 1 0.1 -
A B C D E F G
TSIII
A -
B 1 -
C 1 1 -
D 0.1 0.7 0.1 -
E 0.1 1 0.5 0.1 -
F 0.7 0.7 1 0.1 0.1 -
G 0.1 0.1 0.1 1 0.1 0.1 -
A B C D E F G
TSIV
A -
B 1 -
C 1 1 -
D 0.1 0.1 0.1 -
E 0.1 1 0.5 0.1 -
F 0.5 0.5 1 0.1 0.1 -
G 0.1 0.1 0.1 1 0.1 0.1 -
A B C D E F G
TS I
TS II
TS III
TS IV
PA
EOCENE
OLIGOCENE
MIOCENE
PLI
P

Fig S9

ANNEXES OF CHAPTER 5

SUPLEMENTARY MATERIAL

Table S1**.** GEOSSE: Maximum likelihood and Bayesian parameters (mean and standard deviation, SD) under the full GEOSSE model selected in the analysis of range-dependent diversification rates for the *Hypericum* dated phylogeny in Figure 2. Abbreviations: EP = Eastern Palearctic, WP = Western Palearctic, ITH = Irano-Turanian-Himalayan region, AF = Africa, NE = Nearctic.

		sA mean (SD)	**sB mean (SD)**	**sAB mean (SD)**	**xA mean (SD)**	**xB mean (SD)**	**dA mean (SD)**	**dB mean (SD)**
EP	ML	0.762	0.344	1.532	0.669	0.201	0.090	0.016
	Bayes	1.007 (0.568)	0.352 (0.075)	2.534 (1.956)	0.979 (0.624)	0.220 (0.094)	0.145 (0.097)	0.030 (0.023)
WP	ML	0.150	0.731	1.036	0.000	0.607	0.119	0.002
	Bayes	0.187 (0.038)	0.794 (0.174)	1.314 (0.962)	0.058 (0.046)	0.675 (0.192)	0.129 (0.048)	0.003 (0.002)
ITH	ML	0.150	0.279	12.101	0.000	0.287	0.765	0.000
	Bayes	0.335 (0.280)	0.344 (0.088)	6.14 (2.95)	0.406 (0.451)	0.276 (0.103)	0.506 (0.263)	0.043 (0.057)
AF	ML	0.299	0.426	5.219	0.265	0.268	0.095	0.008
	Bayes	0.398 (0.175)	0.427 (0.092)	3.38 (2.29)	0.383 (0.205)	0.276 (0.111)	0.127 (0.096)	0.011 (0.005)
NE	ML	0.238	0.475	0.240	0.100	0.312	0.020	0.002
	Bayes	0.302 (0.108)	0.508 (0.111)	1.415 (1.499)	0.179 (0.126)	0.35 (0.131)	0.0324 (0.02)	0.003 (0.002)

Table S2. Ancestral niche reconstruction: present values of every major phylogenetic clade for the selected variables.

	Annual Prec max	Annual Prec min	Annual Var Prec max	Annual Var Prec min	Aridity max	Aridity min	Continentality max	Continentality min	Max Month Prec max	Max Month Prec min	Min Month Prec max	Min Month Prec min	Min Month Temp max	Min Month Temp min
Adenotrias	998	245	122	25	22	5	24	11	148	34	53	9	10	-2
Triadenum	1784	457	261	37	34	9	48	17	314	62	60	3	11	-22
Myriandra	1632	358	372	29	29	7	38	2	383	53	62	1	26	-20
Brathys	3615	113	384	18	65	2	36	1	422	21	148	1	25	-16
Androsaemum	1853	21	340	6	36	0	31	11	366	6	74	0	19	-5
Campylosporus	2725	83	301	17	46	1	15	3	349	18	66	0	25	8
Ascyreia	4130	84	384	20	70	1	56	1	386	21	301	0	27	-31
Hirtella	1362	115	337	13	32	3	48	11	346	17	59	0	13	-26
Hypericum	2348	13	384	5	43	0	49	2	386	5	105	0	24	-29
Cratoxyleae	4249	679	380	86	72	12	17	1	453	116	268	1	27	8
Harungana	2725	271	421	72	46	5	12	2	422	77	83	0	25	12
Psorospermum	2725	468	421	96	46	8	12	2	422	98	83	0	25	8
Vismia	3853	206	444	71	65	4	10	1	452	71	164	0	27	6

Table S3. Ancestral niche reconstruction: correlation coefficients for the studied variables.

	Means	Std. Dev.	Ann Prec	Ann Var Prec	Aridity	Continentality	Max Mon Prec	Min Mont Prec	Min Mont Temp
Ann Prec	1411.11	192.46	1.00	0.86	1.00	-0.92	0.89	0.65	0.94
Ann Var Prec	193.19	25.79	0.86	1.00	0.85	-0.96	0.98	0.24	0.95
Aridity	25.25	2.64	1.00	0.85	1.00	-0.91	0.90	0.65	0.93
Continentality	18.00	5.37	-0.92	-0.96	-0.91	1.00	-0.97	-0.33	-1.00
Max Mont Prec	208.14	23.14	0.89	0.98	0.90	-0.97	1.00	0.27	0.96
Min Mont Prec	58.34	8.03	0.65	0.24	0.65	-0.33	0.27	1.00	0.39
Min Mont Temp	6.45	5.18	0.94	0.95	0.93	-1.00	0.96	0.39	1.00

Other articles published during the PhD work

Amazonia Through Time: Andean Uplift, Climate Change, Landscape Evolution, and Biodiversity

C. Hoorn, *et al.*

Science **330**, 927 (2010);
DOI: 10.1126/science.1194585

Amazonia Through Time: Andean Uplift, Climate Change, Landscape Evolution, and Biodiversity

C. Hoorn,[1]* F. P. Wesselingh,[2] H. ter Steege,[3] M. A. Bermudez,[4] A. Mora,[5] J. Sevink,[1] I. Sanmartín,[6] A. Sanchez-Meseguer,[6] C. L. Anderson,[6] J. P. Figueiredo,[7] C. Jaramillo,[8] D. Riff,[9] F. R. Negri,[10] H. Hooghiemstra,[1] J. Lundberg,[11] T. Stadler,[12] T. Särkinen,[13] A. Antonelli[14]*†

The Amazonian rainforest is arguably the most species-rich terrestrial ecosystem in the world, yet the timing of the origin and evolutionary causes of this diversity are a matter of debate. We review the geologic and phylogenetic evidence from Amazonia and compare it with uplift records from the Andes. This uplift and its effect on regional climate fundamentally changed the Amazonian landscape by reconfiguring drainage patterns and creating a vast influx of sediments into the basin. On this "Andean" substrate, a region-wide edaphic mosaic developed that became extremely rich in species, particularly in Western Amazonia. We show that Andean uplift was crucial for the evolution of Amazonian landscapes and ecosystems, and that current biodiversity patterns are rooted deep in the pre-Quaternary.

Pleistocene forest remnants ("refugia") were long held to be responsible for Amazonian diversity (*1*). In the 1990s the centers of diversity, postulated as prime evidence for the refuge theory, were shown to be sampling artifacts (*2*). Over time, the theory was abandoned and an older origin for the Amazonian diversity was proposed (*3*). Perhaps more important, regional diversification events, as inferred from the fossil record and molecular phylogenetic studies, mostly predate the Pleistocene (*4*, *5*). Although the mechanisms of diversification remain elusive and speciation may occur with barriers (*6*) and even without clear barriers (*7*), it is now generally acknowledged that the development of Amazonian biota has been a long and complex process (*3*, *8*).

At the global scale, the Neogene (the 20 million years that preceded the Pleistocene) was a defining period during which much of the present geography and biotic composition was formed (*9*). The process of species diversification is strongly linked to tectonism and climate, both in the terrestrial (*10*, *11*) and marine realms (*12*). The dynamic geologic history of South America should thus be very relevant for understanding the origins of the present diversity.

Recent advances in the fields of Andean and Amazonian geology and phylogenetics have proceeded in parallel. The geosciences community provided new data on mountain building in the Andes and on the timing and types of biotic and paleoenvironmental changes in lowland Amazonia. Climatologists modeled the atmospheric patterns that resulted from the formation of the Andean orographic barrier. At the same time, new molecular analyses based on DNA sequence variation of living organisms shed further light on the sequence and approximate timing of diversifications.

These new data made it clear that the Cenozoic uplift history of the Andes and its effect on regional climate (*13*, *14*) has had a large impact on the landscape evolution in entire northern South America, including Amazonia (*15*, *16*). Although links between the Andean orogeny and neotropical diversification have long been suggested (*17*), only recently have researchers started to explore dated phylogenetic trees [e.g., (*18*, *19*)], in combination with more realistic, complex geological scenarios (*8*, *20*).

Here, we review the timing and extent of mountain building in northern South America and compare it with geologic evidence from sedimentary basins in Amazonia. We explore the origins of Amazonian ecosystems and biodiversity with the use of a combination of geologic (including paleontologic) and ecologic data sets as well as dated molecular phylogenies. Through schematic representation of these findings, we summarize the geologic evolution of this area, outline the age structure of its biodiversity, and provide a guideline for future integrated geologic, biogeographic, and conservation studies.

Amazonia Prior to Andean Influence: An Ancient, River-Dominated Landscape

The area known today as Amazonia was once part of a much larger "pan-Amazonian" region, which, before the late Miocene [until 10 million years ago (Ma)], included the area of the present Amazon, Orinoco, and Magdalena drainage basins (Fig. 1A). At times this region extended to the south, into the northern Paraná region (*21*). We call this vast area pan-Amazonia because we know from the fossil record that a diverse fauna existed, elements of which are now restricted to Amazonia.

Most of Amazonia's geologic history was centered on the Amazon Craton, the hard rock core in the eastern part of South America, but this situation changed during the course of the Cenozoic. Following continental breakup (135 to 100 Ma), both the growing Atlantic Ocean and plate tectonic adjustments along the Pacific margin (*22*) caused deformation within the Amazon Craton, and later the formation of the Andes (figs. S1 to S4) (*23*). Archives of this regional history are stored within a series of north-south–trending foreland basins along the Andes, in the east-west–trending intracratonic basins, and in the Amazon submarine fan in the Atlantic (*24*–*26*).

Testimony to the post-breakup changes on the craton are alluvial and braided river deposits of Cretaceous age that accumulated in the east-west–stretching sedimentary basins. These drainage systems were captured in a "reversed" trunk river with westward flow (*27*), quite dissimilar from the present Amazon River. The drainage divide was initially situated in eastern Amazonia, but during Paleogene times (~65 to 23 Ma) it migrated westward (*25*, *28*), giving way to the precursor of the modern lower Amazon River (Fig. 1, A and B). Toward the end of the Paleogene, the continental divide was located in Central Amazonia and separated east- and west-flowing Amazonian rivers (*24*).

During the Paleogene, the western and northwestern parts of the pan-Amazonian lowlands were characterized by alternating fluvial conditions and marginal marine embayments (*26*). Fossils show that a diverse mammalian fauna including rodents, marsupials, ungulates, and xenarthrans existed in the central-western part of pan-Amazonia [e.g., (*29*)]. Paleogene fossils also reveal diversification of a variety of freshwater catfishes, characins, and cichlids now prominent in Amazonian waters

[1]Paleoecology and Landscape Ecology, Institute for Biodiversity and Ecosystem Dynamics (IBED), University of Amsterdam, Science Park 904, 1098 XH Amsterdam, Netherlands. [2]Nederlands Centrum voor Biodiversiteit Naturalis, P.O. Box 9517, 2300 RA Leiden, Netherlands. [3]Institute of Environmental Biology, Department of Biology, Faculty of Science, Utrecht University, Padualaan 8, 3584 CH Utrecht, Netherlands. [4]Laboratorios de Termocronología y Geomatemáticas, Escuela de Geología, Minas y Geofísica. Facultad de Ingeniería, Universidad Central de Venezuela, Postal Code 1053, Caracas, Venezuela. [5]ECOPETROL, Instituto Colombiano del Petroleo, Piedecuesta, Santander, Colombia. [6]Real Jardin Botanico, CSIC, Plaza de Murillo 2, 28014 Madrid, Spain. [7]Petroleo Brasileiro SA (Petrobras), Av. Republica do Chile, 330, 14o Andar, CEP 20.031-170, Rio de Janeiro, Brazil. [8]Smithsonian Tropical Research Institute, Box 0843-03092, Balboa, Republic of Panama. [9]Instituto de Biologia, Universidade Federal de Uberlândia, Campus Umuarama, Bloco 2D-sala 28, Rua Ceará s/n, Bairro Umuarama, Uberlândia, CEP 38400-902, Minas Gerais, Brazil. [10]Laboratório de Paleontologia, Campus Floresta, Universidade Federal do Acre, Estrada do Canela Fina, Km 12, Cruzeiro do Sul, Acre, CEP 69980-000, AC, Brazil. [11]Department of Ichthyology, Academy of Natural Sciences, 1900 Benjamin Franklin Parkway, Philadelphia, PA 19103, USA. [12]Institute of Integrative Biology, ETH Zürich, Universitätsstrasse 16, 8092 Zürich, Switzerland. [13]Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK. [14]Institute of Systematic Botany, University of Zürich, Zollikerstrasse 107, CH 8008 Zürich, Switzerland.

*To whom correspondence should be addressed. E-mail: carina.hoorn@milne.cc (C.H.); alexandre.antonelli@vgregion.se (A.A.)

†Present address: Gothenburg Botanical Garden, Carl Skottsbergs Gata 22A, 413 19 Göteborg, Sweden, and Department of Plant and Environmental Sciences, University of Gothenburg, Carl Skottsbergs Gata 22B, 413 19 Göteborg, Sweden.

(*21*, *30*). Typical South American mammals such as the xenarthrans (sloths, armadillos, and anteaters), as well as podocnemidid turtles and plant groups such as *Nothofagus*, *Araucaria*, *Gunnera*, and Winteraceae, may have colonized South America through the southern "Gondwanan" connection with Antarctica and Australia, which lasted until the Late Eocene (*31*–*33*). But the role of dispersal versus vicariance in shaping disjunct distributions in the southern hemisphere is intensely debated. Despite continental isolation to the north lasting until the Pliocene, waves of immigrants (e.g., bats and plant families such as Malpighiaceae, Fabaceae, Annonaceae, and Rubiaceae) arrived from the boreotropical regions while caviomorph rodents and platyrrhine primates possibly crossed the Atlantic from Africa (Fig. 2A).

Andean Uplift, a Major Driver for Change in the Amazonian Landscape and Biota

Uplift in the Central and Northern Andes was a partially synchronous process caused by plate tectonic readjustments [(*23*); see also references in (*16*)]. Plate subduction along the Pacific margin caused uplift in the Central Andes during the Paleogene [65 to 34 Ma; see references in (*14*, *16*)]. Posterior plate breakup in the Pacific (~23 Ma) and subsequent collision of the new plates with the South American and Caribbean plates resulted in intensified mountain building in the Northern Andes (figs. S1 to S4) (*16*). Mountain building first peaked in this region by the late Oligocene to early Miocene (~23 Ma), at an age

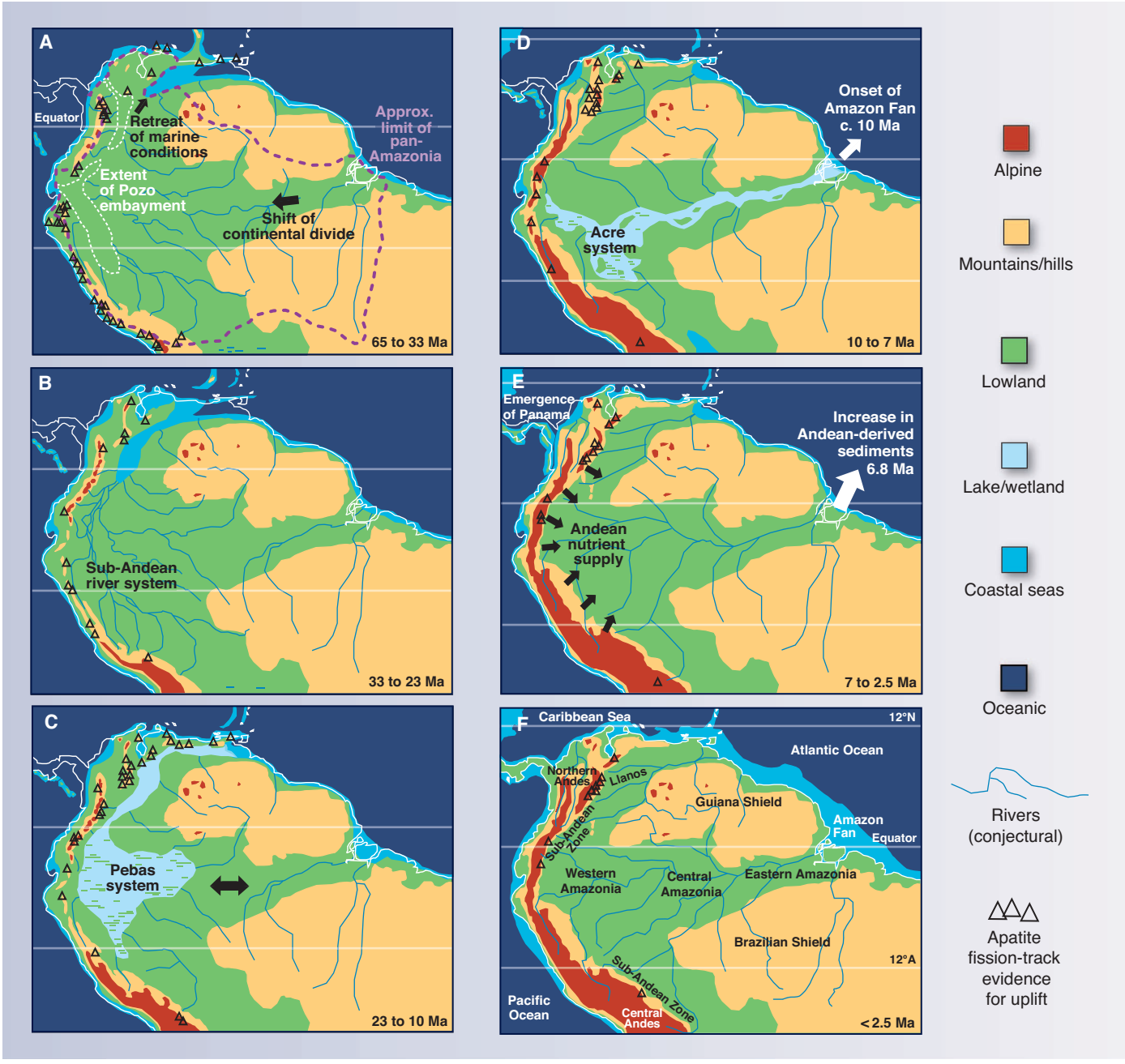


Fig. 1. Paleogeographic maps of the transition from "cratonic" (**A** and **B**) to "Andean"-dominated landscapes (**C** to **F**). (A) Amazonia once extended over most of northern South America. Breakup of the Pacific plates changed the geography and the Andes started uplifting. (B) The Andes continued to rise with the main drainage toward the northwest. (C) Mountain building in the Central and Northern Andes (~12 Ma) and wetland progradation into Western Amazonia. (D) Uplift of the Northern Andes restricted "pan-Amazonia" and facilitated allopatric speciation and extirpation [e.g., (*21*)]. (E) The megawetland disappeared and *terra firme* rainforests expanded; closing of Panama Isthmus and start of GABI. (F) Quaternary. Note that South America migrated northward during the course of the Paleogene.

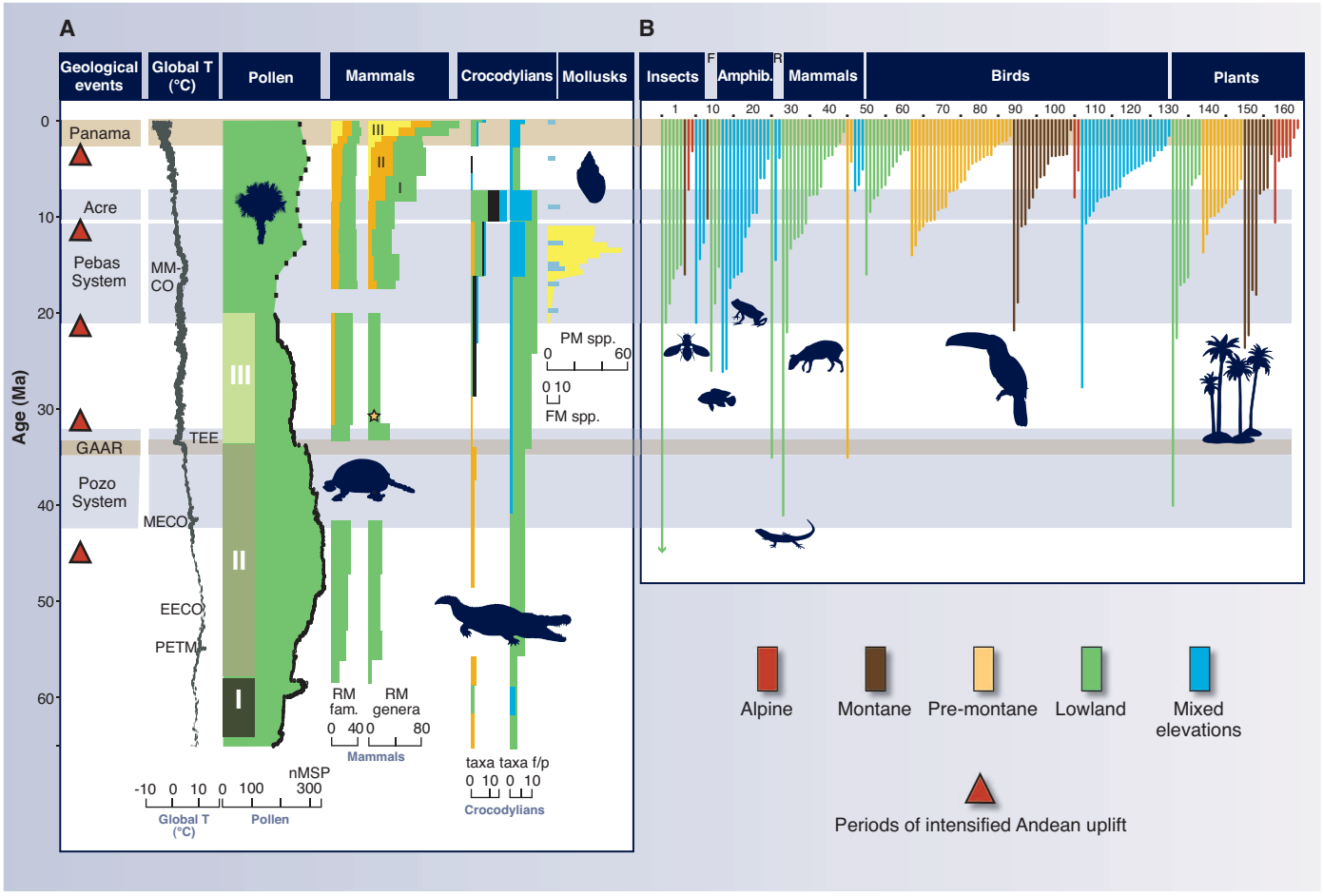


Fig. 2. Biotic changes in Amazonia through time (*23*). (**A**) The Cenozoic fossil record of the tropical lowlands reveals the timing of biotic turnover. Paleogene floral diversity (from pollen records) increased with high temperature, but in the Neogene it was unrelated and remained relatively high even under cooler conditions. Mollusks and crocodiles diversified with the onset of the Miocene megawetlands and declined with its demise. The fossil record, as is shown for the caimanine crocodiles (blue in the right column), is nonetheless incomplete when compared to minimum expected numbers of species (green in the right column) derived from phylogenetic reconstructions (*23*). Late Neogene mammal diversification was particularly strong among North American derived taxa. MMCO, Middle Miocene Climate Optimum; PETM, Paleocene-Eocene Thermal Maximum; MECO, Middle Eocene Climate Optimum; EECO, Early Eocene Climate Optimum; TEE, Terminal Eocene Event; GAAR, Greater Antilles-Aves Ridge; nMSP, number of pollen morphospecies; RM, running mean; f/p, from the fossil record or as based on caimanine phylogeny; FM, fluvial mollusk; PM, Pebasian endemic mollusk species. Crocodylians: Left column, number of species from fossil record; right column, number of caimanine species from fossil record versus number of lineages (orange, non-eusuchian crocodyliforms; green, Caimaninae; black, Gavialoidea; blue, Crocodylidae). Global temperature curve is based on (*68*). Abbreviations are further explained in (*23*). (**B**) Diversification of modern lineages revealed from molecular phylogenies. The lines illustrate the approximate timing of diversification for genera of animals and plants in northern South America, in relation to the elevation zone they inhabit (lowland, 0 to 500 m; premontane, 500 to 1500 m; montane, 1500 to 3000 m; alpine, 3000 to 4800 m). Nearly all living genera in northern South America have a pre-Quaternary origin, but ages of taxa differ between major elevation zones. Several highland genera are fairly young; lowland genera are a mixture of young and old lineages. Numbers above individual lines refer to table S1, where additional details are given.

that coincides with the diversification of the first modern montane plant and animal genera (Fig. 2B). However, the most intense peaks of Andean mountain building followed during the late middle Miocene (~12 Ma, Fig. 1C) and early Pliocene (~4.5 Ma, Fig. 1E and figs. S3 to S5) (*16*). Plate reorganization ultimately resulted in closing of the Panama Isthmus during the Pliocene (at ~3.5 Ma) (*34*) and led to the Great American Biotic Interchange (see below).

Mountain building in the Andes generated tectonic load and renewed accommodation space in the adjacent foreland basins. As mountain building progressed and a critical elevation (~2000 m; figs. S3 to S5) was surpassed, rainfall increased along the eastern flank. This coupling of tectonic and climatic processes resulted in further uplift, erosion, and water and sediment supply (*13*, *14*, *35*) and is in accordance with changes in the depositional record of the Andean foreland and Amazonia (fig. S5). However, the Andean sediment flux that engulfed lowland Amazonia (*36*) was not continuous; intramontane basins and perimontane basins may have captured influx for periods of millions of years, resulting in pulses of deposition eastward.

Parallel to intensified uplift in the Andes, a large wetland of shallow lakes and swamps developed in Western Amazonia (Fig. 1C) (*37*). These new aquatic environments of the "Pebas" system were colonized by rapidly radiating endemic invertebrate faunas composed of mollusks and ostracods (*38*). This was also the stage for a diverse reptile fauna including gharials, caimans, and turtles (Fig. 2A). One of the most remarkable representatives of this now-extinct fauna was *Purussaurus,* the largest known caiman, which reached ~12 m in length (*39*).

The wetland fragmented the preexisting rainforests, yet a diverse forest that already bore resemblance to the modern forest (in terms of plant family composition) remained at the margins of this new aquatic system (*15*, *40*). Although lower than in the Paleogene, plant diversity (as indicated by pollen types) peaked at 13 Ma, near the end of

the Middle Miocene Climatic Optimum (Fig. 2A). Geochemical evidence from mollusk shells further indicates that a modern type of monsoonal climate was already present and provided a seasonal water influx into the wetland system (*41*). Terrestrial taxa such as xenarthrans, *Gonatodes* geckos, and leaf beetles, as well as cichlid fish in the aquatic environments, lived and diversified in the wetlands (Fig. 2B and table S1).

Taxa of marine ancestry in the Miocene (*42*) or earlier (*43*), such as potamotrygonid stingrays, thrived in the Amazonian freshwater wetlands. Periods with somewhat elevated salinities are also indicated by benthic foraminifera, barnacles, (marginal) marine mollusks, and the geochemical signature in the mollusk shells (*44*). These marine invertebrates, however, were Neogene arrivals and disappeared with the withdrawal of marginal marine conditions. Other indicators of marine influence in the wetlands were dinoflagellates, pollen from mangrove trees, and marine ichnofossils. Biogeographic reconstructions based on phylogenies also fit this scenario (*8*, *20*, *42*). Despite such evidence, the extent of marine influence in Amazonia is still debated (*45*).

By the end of the middle Miocene (~12 Ma), faster and more widespread Andean mountain building prompted peak topographic growth. This created deep canyon incision and erosion in the Central and Northern Andes, especially in the Eastern Cordilleras and in the Venezuelan Andes (figs. S1 to S4) (*16*, *46*), where alluvial megafans developed (*47*, *48*). It also coincided with raised sedimentation rates in the Andean foreland basins that eventually became overfilled. At ~10 Ma, coinciding with global sea level drop and climate cooling, Andean sediments reached the Atlantic coast through the Amazon drainage system, and the Amazon River became fully established at ~7 Ma (*24*, *49*).

Meanwhile, the Western Amazonian wetland changed from a lacustrine to a fluvial or fluviotidal system (Fig. 1D) (*37*, *45*, *50*), which resembled the present-day Pantanal in southern Amazonia (*45*). This so-called "Acre" system harbored a very rich aquatic vertebrate fauna that included mega-sized gharials, caimanines, and side-neck turtles (*39*), which eventually declined with the disappearance of megawetlands in Western Amazonia at ~7 Ma (Fig. 2A) (*21*, *38*, *39*). Most of the endemic mollusk fauna was unable to adapt to the initial fluvial conditions and was strongly reduced around 10 Ma (*38*). The floodplains of this system were dominated by grasses (*51*) and were inhabited by a more diverse xenarthran fauna than at present (*52*).

Preliminary palynological evidence indicates a ~10 to 15% increase of plant diversity between ~7 and 5 Ma, shortly after the wetlands were replaced by forested habitats (Fig. 2A). Molecular studies of tree genera such as *Guatteria* (Annonaceae, ~250 species) and *Inga* (Fabaceae, ~300 species) show a similar trend of rapid diversification following the demise of Amazonian wetlands (*53*, *54*). This suggests that the establishment of terrestrial conditions in Western Amazonia may have been an important prerequisite for the diversification of the current biota of this region. However, the actual triggers of speciation in these and other cases may have been much more complex, involving factors such as soil adaptation and plant-herbivore interactions (*55*).

Western Amazonia from then on bore the key geographic features of the landscape as we know it today (Fig. 1, E and F). It had changed from a drowning, negative relief into a positive relief incised by an increasingly entrenched river system with high sediment load. By the late Miocene, good swimmers such as proboscideans had crossed the relatively small seaway that remained between Central and South America and were at the forefront of a major immigration wave (*56*, *57*).

The final scenes of this history are characterized by further Andean uplift (Fig. 1F), closure of the Panama Isthmus (~3.5 Ma), the Quaternary ice ages (2.5 to 0.01 Ma), and restriction of megafans in the foreland basin zone. This, together with neotectonic processes in Amazonian lowlands (*28*), caused uplift of the Neogene deposits, development of widespread river terrace systems, and readjustments of river patterns, and led to the mosaic-type landscape of the present (*58*). The accelerated uplift phases during the last 10 Ma fostered spectacular radiations of highland plants such as lupines (*59*), as well as tanagers, bumblebees, and some rodents (Fig. 2B and table S1). This was also a time of extensive migration, when both Amazonia and the new montane habitats in the Andes were colonized by taxa of North American descent during the Great American Biotic Interchange (GABI) (*56*).

The GABI caused decline in the number of endemic South American mammal families during the Pliocene and especially the Quaternary. However, the overall generic diversity of South American mammal taxa remained stable, and the total number of genera increased by the strong diversification of taxa derived from North American immigrants (*56*) (Fig. 2A). Molecular studies suggest that many bird lineages also took part in the GABI (*60*, *61*). By contrast, plants have been more capable of overseas dispersal, and many lineages crossed the Panama Isthmus before its final closure (*62*), whereas others probably reached South America directly from Africa (*63*). These results, based on molecular and fossil studies, suggest that immigrants from other landmasses have played an important role in the historic assembly of the Amazonian biota (*64*).

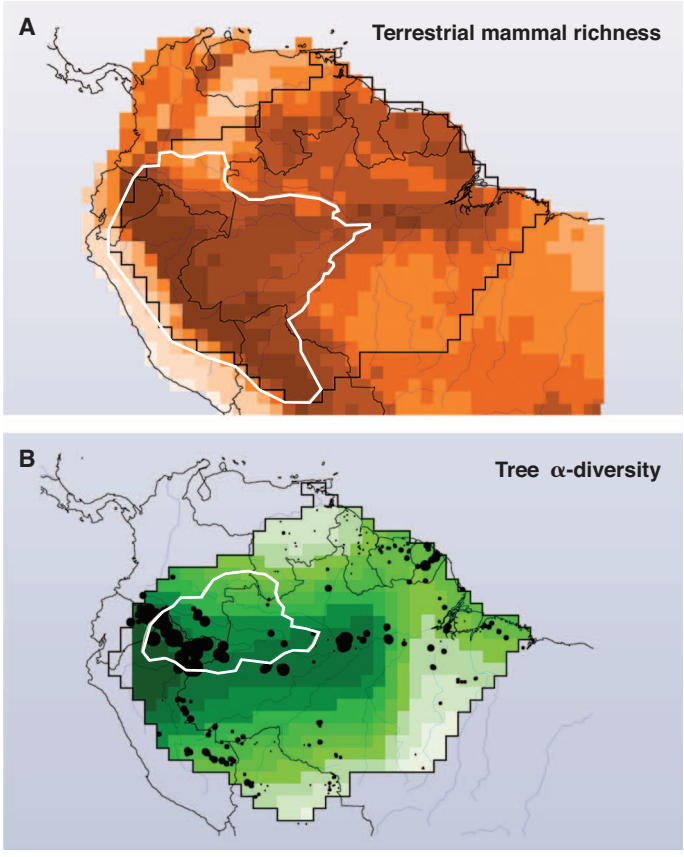


Fig. 3. Present Amazonian diversity patterns. See figs. S6 and S7 for depictions of the close relationship among Amazonian geology, soils, climate, and diversity. (**A**) Terrestrial mammal richness (range: lightest color, 2 to 10 species; darkest, 89 to 109 species) (*69*); white polygon denotes relatively rich soils (fig. S6C). (**B**) Tree α-diversity (*66*). Black dots: local tree α-diversity on 1-ha plots (n = 752); Fisher's α ranges from 3.6 to 300; green shades: loess spatial interpolation of 1-ha values (6 to 117); white polygon: area of least severe water shortage (see fig. S6D).

Can Geologic History Help Us Understand Present Biodiversity in Amazonia?

A comparison of present biodiversity patterns with geologic and edaphic units shows that the highest concentrations of terrestrial mammal and amphibian richness are found on Western Amazonian soils that developed on the Neogene (Andean) sediments (Fig. 3A and figs. S6 and S7). These soils show much higher variation in levels of nutrients and are in stark contrast to generally nutrient-poor soils on the craton in Eastern Amazonia (*65*). Forest productivity and forest dy-

namics are also higher on these soils (fig. S8), which suggests that bedrock composition, diversity, and ecosystem productivity are interrelated (*66*).

Water geochemistry, sediment composition, and fertility of floodplains further confirm the disproportionate richness in nutrients of the Andean system versus the relative nutrient poverty in the "cratonic" aquatic system (*67*). It seems paradoxical that the old Amazon Craton, which had the opportunity to accumulate taxa for a much longer period than the young areas in Western Amazonia, has fewer species, genera, and families.

Nutrients and habitat heterogeneity are paramount in Amazonian diversity, but they are not the only ingredient. Tree α-diversity (i.e., the diversity measured on 1-ha plots) peaks in the wetter, less seasonal part of Western Amazonia (Fig. 3B), which suggests a role for climate in sustaining (and perhaps also driving) diversity (*66*). By contrast, the highest levels of mammal diversity appear little affected by rainfall seasonality, from aseasonal Ecuador down to highly seasonal Bolivia (Fig. 3A and fig. S6D); this suggests that additional factors such as productivity need to be considered.

Although the transition from a "cratonic" to an "Andean"-dominated system was a fundamental change in the evolution of Amazonian landscapes and species composition, all data suggest that this switch was a complex, stepwise process. Species accumulation was driven by more than one single, overarching mechanism, and Amazonian biodiversity was certainly not a by-product of just Pleistocene ice ages, but resulted from a much more extended period of evolution. However, after the draining of the wetlands (late Miocene), diversification in Western Amazonia must have been particularly rapid, as the diversity of this area greatly outnumbers the diversity in the cratonic areas.

Many outstanding research questions concerning Amazonia remain. Understanding the mechanisms that underlie the assembly and evolution of Amazonian biodiversity continues to be a major challenge that will require hitherto unrealized interdisciplinary scientific collaboration. Evolutionary studies linked to molecular phylogenies and fossil assemblages should focus on Neogene records and on species-rich but poorly sampled areas. Future research should be concentrated on the interface between the Cenozoic and cratonic areas, and on the transition zone between the Andes and Western (lowland) Amazonia (fig. S6). This area, together with the southern fringe of Amazonia, has become rapidly occupied by humans but nonetheless remains scientifically poorly known.

References and Notes

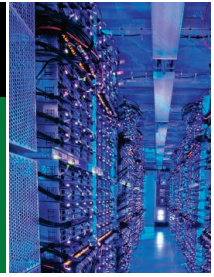
1. J. Haffer, *Science* **165**, 131 (1969).
2. B. W. Nelson, C. A. C. Ferreira, M. F. da Silva, M. L. Kawasaki, *Nature* **345**, 714 (1990).
3. M. B. Bush, *J. Biogeogr.* **21**, 5 (1994).
4. C. Jaramillo, M. J. Rueda, G. Mora, *Science* **311**, 1893 (2006).
5. V. Rull, *Mol. Ecol.* **17**, 2722 (2008).
6. F. E. Hayes, J. A. N. Sewlal, *J. Biogeogr.* **31**, 1809 (2004).
7. M. A. M. de Aguiar, M. Baranger, E. M. Baptestini, L. Kaufman, Y. Bar-Yam, *Nature* **460**, 384 (2009).
8. J. C. Santos *et al.*, *PLoS Biol.* **460**, e1000056 (2009).
9. P. E. Potter, P. Szatmari, *Earth Sci. Rev.* **96**, 279 (2009).
10. M. J. Kohn, T. J. Fremd, *Geology* **36**, 783 (2008).
11. J. A. Finarelli, C. Badgley, *Proc. Biol. Sci.* **277**, 2721 (2010).
12. W. Renema *et al.*, *Science* **321**, 654 (2008).
13. N. Insel, C. J. Poulsen, T. A. Ehlers, *Clim. Dyn.* (2009).
14. C. J. Poulsen, T. A. Ehlers, N. Insel, *Science* **328**, 490 (2010); 10.1126/science.1185078.
15. C. Hoorn, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **105**, 267 (1993).
16. A. Mora *et al.*, in *Amazonia, Landscape and Species Evolution*, C. Hoorn, F. P. Wesselingh, Eds. (Wiley, Oxford, 2010), pp. 38–60.
17. A. H. Gentry, *Ann. Mo. Bot. Gard.* **69**, 557 (1982).
18. J. S. Albert, N. R. Lovejoy, W. G. R. Crampton, *J. S. Am. Earth Sci.* **21**, 14 (2006).
19. R. T. Brumfield, S. V. Edwards, *Evolution* **61**, 346 (2007).
20. A. Antonelli, J. A. A. Nylander, C. Persson, I. Sanmartín, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 9749 (2009).
21. J. G. Lundberg *et al.*, in *Phylogeny and Classification of Neotropical Fishes*, M. Malabarba, R. E. Reis, R. P. Vari, Z. M. Lucena, C. A. S. Lucena, Eds. (Edipucrs, Porto Alegre, Brazil, 1998), pp. 13–48.
22. B. L. Isacks, *J. Geophys. Res.* **93**, 3211 (1988).
23. See supporting material on *Science* Online.
24. J. Figueiredo, C. Hoorn, P. van der Ven, E. Soares, *Geology* **37**, 619 (2009).
25. J. R. Wanderley-Filho, J. F. Eiras, P. R. da Cruz-Cunha, P. H. van der Ven, in *Amazonia, Landscape and Species Evolution*, C. Hoorn, F. P. Wesselingh, Eds. (Wiley, Oxford, 2010), pp. 29–37.
26. M. Roddaz *et al.*, in *Amazonia, Landscape and Species Evolution*, C. Hoorn, F. P. Wesselingh, Eds. (Wiley, Oxford, 2010), pp. 61–88.
27. R. W. Mapes, thesis, University of North Carolina at Chapel Hill (2009).
28. J. B. Sena Costa, R. Léa Bemerguy, Y. Hasui, M. da Silva Borges, *J. S. Am. Earth Sci.* **14**, 335 (2001).
29. K. E. Campbell Jr., Ed., *The Paleogene Mammalian Fauna of Santa Rosa, Amazonian Peru* (Natural History Museum of Los Angeles County, Los Angeles, 2004).
30. M. C. Malabarba, L. R. Malabarba, C. Del Papa, *J. Vertebr. Paleontol.* **30**, 341 (2010).
31. S. F. Vizcaíno, G. J. Scillato-Yané, *Antarct. Sci.* **7**, 407 (1995).
32. I. Sanmartín, F. Ronquist, *Syst. Biol.* **53**, 216 (2004).
33. B. P. Noonan, P. T. Chippindale, *Am. Nat.* **168**, 730 (2006).
34. A. G. Coates *et al.*, *Geol. Soc. Am. Bull.* **104**, 814 (1992).
35. M. R. Strecker *et al.*, *Geology* **37**, 643 (2009).
36. R. Aalto, T. Dunne, J.-L. Guyot, *J. Geol.* **114**, 85 (2006).
37. C. Hoorn, F. P. Wesselingh, J. Hovikoski, J. Guerrero, in *Amazonia, Landscape and Species Evolution*, C. Hoorn, F. P. Wesselingh, Eds. (Wiley, Oxford, 2010), pp. 123–142.
38. F. P. Wesselingh, J. Salo, *Scr. Geol.* **133**, 439 (2006).
39. D. Riff, P. S. R. Romano, G. R. Oliveira, O. A. Aguilera, in *Amazonia, Landscape and Species Evolution*, C. Hoorn, F. P. Wesselingh, Eds. (Wiley, Oxford, 2010), pp. 259–280.
40. D. Pons, D. De Franceschi, *Bull. Geosci.* **82**, 343 (2007).
41. R. J. G. Kaandorp *et al.*, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **221**, 1 (2005).
42. N. R. Lovejoy, J. S. Albert, W. G. R. Crampton, *J. S. Am. Earth Sci.* **21**, 5 (2006).
43. M. R. De Carvalho, J. G. Maisey, L. Grande, *Bull. Am. Mus. Nat. Hist.* **284**, 1 (2004).
44. H. B. Vonhof *et al.*, *Geol. Soc. Am. Bull.* **115**, 983 (2003).
45. E. M. Latrubesse *et al.*, *Earth Sci. Rev.* **99**, 99 (2010).
46. M. Bermúdez *et al.*, *Tectonics* **29**, TC5009 (2010).
47. B. K. Horton, P. G. DeCelles, *Basin Res.* **13**, 43 (2001).
48. C. E. Uba, M. R. Strecker, A. K. Schmitt, *Geology* **35**, 979 (2007).
49. J. Figueiredo, C. Hoorn, P. van der Ven, E. Soares, *Geology* **38**, e213 (2010).
50. J. Hovikoski *et al.*, *Geol. Soc. Am. Bull.* **119**, 1506 (2007).
51. S. A. F. da Silva-Caminha, C. A. Jaramillo, M. L. Absy, *Palaeontogr. Abt. B* **283**, 1 (2010).
52. F. R. Negri, thesis, Pontifícia Universidade Católica do Rio Grande do Sul (2004).
53. R. H. J. Erkens, L. W. Chatrou, J. W. Maas, T. van der Niet, V. Savolainen, *Mol. Phylogenet. Evol.* **44**, 399 (2007).
54. J. E. Richardson, R. T. Pennington, T. D. Pennington, P. M. Hollingsworth, *Science* **293**, 2242 (2001).
55. P. V. A. Fine, D. C. Daly, G. Villa Muñoz, I. Mesones, K. M. Cameron, *Evolution* **59**, 1464 (2005).
56. L. G. Marshall, R. L. Cifelli, *Palaeovertebrata* **19**, 169 (1990).
57. B. J. MacFadden, *Trends Ecol. Evol.* **21**, 157 (2006).
58. T. Toivonen, S. Mäki, R. Kalliola, *J. Biogeogr.* **34**, 1374 (2007).
59. C. Hughes, R. Eastwood, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10334 (2006).
60. B. T. Smith, J. Klicka, *Ecography* **33**, 333 (2010).
61. J. T. Weir, E. Bermingham, D. Schluter, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 21737 (2009).
62. S. Cody, J. E. Richardson, V. Rull, C. Ellis, R. T. Pennington, *Ecography* **33**, 326 (2010).
63. S. Renner, *Int. J. Plant Sci.* **165** (suppl. 4), S23 (2004).
64. R. T. Pennington, C. W. Dick, *Philos. Trans. R. Soc. Ser. B* **359**, 1611 (2004).
65. C. A. Quesada *et al.*, *Biogeosci. Discuss.* **6**, 3923 (2009).
66. H. ter Steege, Amazon Tree Diversity Network, RAINFOR (Amazon Forest Inventory Network), in *Amazonia: Landscape and Species Evolution*, C. Hoorn, F. Wesselingh, Eds. (Wiley, Oxford, 2010), pp. 349–359.
67. M. E. McClain, R. J. Naiman, *Bioscience* **58**, 325 (2008).
68. J. C. Zachos, G. R. Dickens, R. E. Zeebe, *Nature* **451**, 279 (2008).
69. M. F. Tognelli, D. A. Kelt, *Ecography* **27**, 427 (2004).

70. We thank all colleagues who shared their data; M. F. Tognelli for mammal richness data; and B. P. Kohn, M. Bernet, P. van der Beek, R. T. Pennington, S. B. Kroonenberg, B. Bookhagen, C. Uba, and three anonymous reviewers for constructive comments on the manuscript. Supported by the Osk. Huttunen Foundation, the Ella and Georg Ehrnrooth Fund, and the Helsingin Sanomain Saatio (T.S.) and by the Academy of Sciences of Finland (F.P.W.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/330/6006/927/DC1
Materials and Methods
Figs. S1 to S8
Tables S1 to S3
References

10.1126/science.1194585



LETTERS

edited by Jennifer Sills

Recognizing Scientists and Technologists

ON 17 NOVEMBER 2010, PRESIDENT OBAMA PRESENTED THE National Medals of Science and the National Medals of Technology and Innovation. These medals are the highest honor that the nation can bestow in science and technology, yet they are rarely mentioned by the popular media. Because Congress does not appropriate funds to implement the "outreach" of these medals, for many years the only national recognition was a private award ceremony with the President.

In 1991, George Rathmann, one of the founders of the biotech industry, facilitated the formation of what is now the National Science and Technology Medals Foundation. The mission of the Foundation is to promote the National Medal Laureates as role models for students and thereby encourage interest in science and math. To accomplish this goal, the Foundation hosts a banquet in conjunction with the White House ceremony. This banquet features videos highlighting the technical accomplishments of the Laureates, which then become the basis for stories that appear throughout the country.

Over the years, the Foundation has accumulated a wealth of electronic material on the Laureates, including biographies, interviews, and descriptions of their accomplishments (*1*). This recognition not only is a way to recognize the Laureates' enormous efforts, but also serves to focus our attention on the seminal ideas in science, mathematics, and engineering. The stories behind these accomplishments often provide inspiration to others, which is essential to promote further achievements.

ROBERT M. WHITE

Materials Science and Engineering, Stanford University, Stanford, CA 94305, USA. E-mail: RMWhite@stanford.edu

Reference

1. National Science and Technology Medals Foundation (www.nationalmedals.org).

Genetics-Based Field Studies Prioritize Safety

M. ENSERINK'S NEWS OF THE WEEK STORY ON the open release trials of genetically modified mosquitoes in the Cayman Islands ("GM mosquito trial alarms opponents, strains ties in Gates-funded project," 19 November 2010, p. 1030) highlights the growing pains associated with bringing new technologies out of the laboratory into the field. Unlike for vaccines, drugs, and insecticides, no industry-wide standards are yet in place to guide either public or private efforts in the development of these technologies. However, it is important for the public to know that the scientists working on these new technologies are aggressively supporting the formulation of best practices for their safe, efficient, ethical, and regulated application, and are reaching out to experts from a range of relevant disciplines for advice and counsel. A series of publications document the evolution of this process (*1*–*5*). Indeed, efforts are currently under way to develop a guidance framework for quality standards to assess safety and efficacy and to address regulatory, legal, social, and cultural issues, as recommended by an international consultation held at the World Health Organization in 2009 (*5*). Thus, although we have not achieved harmonized international standards, as has taken decades for other technologies, we are much closer than most people realize. We recognize the need to ensure that our enthusiasm for the promise of these approaches as powerful public health tools does not outstrip our responsibility to apply scientifically validated and socially acceptable product development practices. The tragedy would be if this important but complex birthing process were to stifle creativity in the development of not only genetics-based solutions, but all truly novel approaches that seek to reduce the serious health threat of diseases such as malaria and dengue fever. We hope that debates over specific circumstances do not cloud the urgent need for the development and deployment of new tools to mitigate these disease scourges.

ANTHONY A. JAMES

Departments of Microbiology and Molecular Genetics, and Molecular Biology and Biochemistry, University of California, Irvine, CA 92697–3900, USA. E-mail: aajames@uci.edu

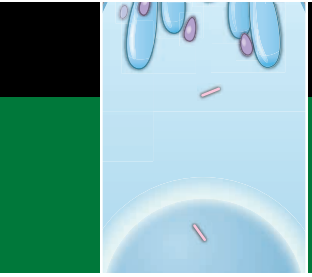
References

1. M. Benedict *et al.*, *Vector-Borne Zoonotic Dis.* **8**, 127 (2008).
2. J. V. Lavery, L. C. Harrington, T. W. Scott, *Am. J. Trop. Med. Hyg.* **79**, 312 (2008).
3. J. V. Lavery *et al.*, *Trends Parasitol.* **26**, 279 (2010).
4. J. Mumford, *Asian Pac. J. Mol. Biol. Biotechnol.* **17**, 91 (2009).
5. WHO/TDR, "Progress and prospects for the use of genetically-modified mosquitoes to inhibit disease transmission," Report on planning meeting 1: Technical consultation on current status and planning for future development of genetically-modified mosquitoes for malaria and dengue control (WHO/TDR publications 10.2471/TDR.10.978-924-1599238, 2010).

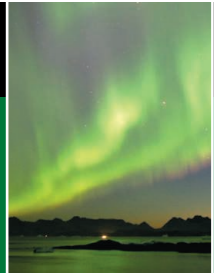
Origins of Biodiversity

THE ORIGIN OF THE HIGH NEOTROPICAL BIODIVERSITY has been a controversial topic since Darwin. The debate has focused on the relative influences of the climate changes during the Pleistocene (the past

Published by AAAS



Parasite replication
409



SPORE Prize Essay
413

~2.6 million years) and the tectonic and geographical reorganizations that occurred before the Pleistocene (*1*). In their Review "Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity" (12 November 2010, p. 927), C. Hoorn *et al*. conclude that the biodiversity patterns of the Amazon basin were largely shaped before the Pleistocene, facilitated by the Andean uplift. The authors dismiss Pleistocene diversification by arguing that, in the Neotropics, the refuge hypothesis (proposing that species diversified in isolated forests during glacial periods) has already been abandoned, and that fossils and molecular phylogenetics support mostly pre-Pleistocene diversification.

However, Pleistocene diversification could have resulted from a variety of mechanisms other than isolated forest remnants (*2*–*4*). Furthermore, Hoorn *et al*. cite my meta-analysis (*5*) in support of the pre-Pleistocene diversification, yet the conclusions I drew from that study contradict those of Hoorn *et al*.'s Review. I concluded that about half of the dated extant neotropical species originated during the Pleistocene and the other half before it, and that speciation proceeded in a continuous fashion with no evident bursts (*5*). In addition, phylogenetic evidence provided by Hoorn *et al*. is based on the dating of complexes of extant species (the crown clades) that in fact records the age of the oldest species within each group (*6*) but not necessarily the age of all the extant species, which should be necessarily younger. This overestimates pre-Pleistocene diversification. Earth's biodiversity gradients are the result of a long and complex history of evolutionary trends, mediated by ecological processes and governed by external forces, in which not only speciation but also extinction should be considered, especially in extratropical areas (*7*).

The topic requires the synergy of many disciplines, in a wide range of spatial and temporal scales. Pleistocene speciation is one more element and should not be neglected; after all, we ourselves are a Pleistocene species barely 200,000 years old.

VALENTÍ RULL

Palynology and Paleoecology Laboratory, Botanical Institute of Barcelona (CSIC-ICUB), 08038 Barcelona, Spain. E-mail: vrull@ibb.csic.es

References

1. M. B. Bush, H. Hooghiemstra, in *Climate Change and Biodiversity*, T. E. Lovejoy, L. Hannah, Eds. (Yale Univ. Press, New Haven, CT, 2005), pp. 125–137.
2. M. B. Bush, *J. Biogeogr*. **21**, 5 (1994).
3. B. P. Noonan, P. Gaucher, *Mol. Ecol*. **14**, 3017 (2005).
4. V. Rull, *J. Biogeogr*. **31**, 1 (2005).
5. V. Rull, *Mol. Ecol*. **17**, 2722 (2008).
6. M. J. Benton, *Biol. Rev*. **75**, 633 (2000).
7. M. S. McGlone, *Glob. Ecol. Biogeogr. Lett*. **5**, 309 (1996).

Response

IN OUR REVIEW, WE LINK THE OUTSTANDING species richness in northern South America to the cataclysmic changes induced by Andean mountain building. Evidence for this is the correlation between sedimentary records, the paleontological record, dated molecular phylogenies, and present species distributions. Our conclusions contradict the hypothesis that has dominated for more than 40 years: that the outstanding levels of Neotropical species richness and current distribution patterns were mainly produced by Quaternary climatic fluctuations (*1*, *2*), i.e., in the past 2.6 million years. All evidence in our meta-analysis points toward an older origin of Amazonian biodiversity.

Rull argues that we ignore Quaternary evidence on speciation, in part by erroneously referring to his previous meta-analysis (*3*) as evidence for pre-Quaternary diversification. Rull's finding that about half of all extant species analyzed originated during Quaternary times (*3*) is not surprising. Assuming the average species longevity is some 100,000 to a couple of million years (*3*–*5*), at any point in time we would expect to find that most species originated in the past few million years. Rull's evidence that extant species originated recently does not contradict the idea that the total number of species was just as high (and for most organism groups higher) before the Quaternary, even if the species that existed then have since become extinct. Moreover, if Pleistocene glaciations had indeed produced most of the species richness observed today—as implied in the original formulation of the "refuge theory" (*1*)—this would unrealistically imply that all previous diversity was produced by entirely different mechanisms. This realization severely undermines the role of glaciation dynamics in accounting for Neotropical species richness.

Rull's suggestion that we overestimated pre-Quaternary diversification by using genera instead of species as taxonomic units in our meta-analysis is misleading. Extinction is more likely to affect older lineages than younger ones—simply because species that have arisen recently have had less time to go extinct (*6*)—meaning that Pre-Quaternary speciation events were probably underestimated in Rull's meta-analysis (*3*). Stochastic diversification models (*6*) can correct for the effect of background extinction in diversification rate estimates, but these models have proven unrealistic because of their oversimplified assumptions (*7*) and sensitivity to incomplete taxon sampling (*8*), a common feature in Neotropical phylogenies. Estimates of crown ages of genera are, arguably, less sensitive to incomplete taxon sampling, because in most species-level phylogenies, sampling is aimed to cover the geographic and morphological variation within a genus. This should lead to more robust age estimation of deeper nodes even when many species are missing.

Angel Falls, Venezuela. Debate continues about when and where neotropical biodiversity developed.

CREDIT: LUIS CARILLO, VENEZUELA

Published by AAAS

The data we assembled show that the blueprint of present Amazonia was laid out in pre-Quaternary times, but they also have the potential to provide us clues on how the rainforest may react to future global warming. It is also clear that Amazonian biota withstood large geodynamic (*9*) and climatic fluctuations but that humans, the young product of Quaternary evolution, pose the biggest threat to this wealth of biodiversity.

C. HOORN,[1] F. P. WESSELINGH,[2] H. TER STEEGE,[3] M. A. BERMUDEZ,[4] A. MORA,[5] J. SEVINK,[1] I. SANMARTÍN,[6] A. SANCHEZ-MESEGUER,[6] C. L. ANDERSON,[6] J. P. FIGUEIREDO,[7] C. JARAMILLO,[8] D. RIFF,[9] F. R. NEGRI,[10] H. HOOGHIEMSTRA,[1] J. LUNDBERG,[11] T. STADLER,[12] T. SÄRKINEN,[13] A. ANTONELLI[14,15]*

[1]Paleoecology and Landscape Ecology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, 1098 XH Amsterdam, Netherlands. [2]Nederlands Centrum voor Biodiversiteit Naturalis, 2300 RA Leiden, Netherlands. [3]Institute of Environmental Biology, Department of Biology, Faculty of Science, Utrecht University, 3584 CH Utrecht, Netherlands. [4]Laboratorios de Termocronología y Geomatemáticas, Escuela de Geología, Minas y Geofísica, Facultad de Ingeniería, Universidad Central de Venezuela, 1053, Caracas, Venezuela. [5]ECOPETROL, Instituto Colombiano del Petroleo, Piedecuesta, Santander, Colombia. [6]Real Jardin Botanico, CSIC, 28014 Madrid, Spain. [7]Petroleo Brasileiro SA (Petrobras), CEP 20.031-170, Rio de Janeiro, Brazil. [8]Smithsonian Tropical Research Institute, Balboa, Republic of Panama. [9]Instituto de Biologia, Universidade Federal de Uberlândia, Campus Umuarama, Bairro Umuarama, Uberlândia, CEP 38400-902, Minas Gerais, Brazil. [10]Laboratório de Paleontologia, Campus Floresta, Universidade Federal do Acre, Estrada do Canela Fina, Cruzeiro do Sul, Acre, CEP 69980-000, AC, Brazil. [11]Department of Ichthyology, Academy of Natural Sciences, Philadelphia, PA 19103, USA. [12]Institute of Integrative Biology, ETH Zürich, 8092 Zürich, Switzerland. [13]Department of Plant Sciences, University of Oxford, Oxford OX1 3RB, UK. [14]Gothenburg Botanical Garden, Carl Skottsbergs Gata 22A, 413 19 Göteborg, Sweden. [15]Department of Plant and Environmental Sciences, University of Gothenburg, Carl Skottsbergs Gata 22B, 413 19 Göteborg, Sweden.

*To whom correspondence should be addressed. E-mail: alexandre.antonelli@vgregion.se

Letters to the Editor

Letters (~300 words) discuss material published in *Science* in the past 3 months or matters of general interest. Letters are not acknowledged upon receipt. Whether published in full or in part, Letters are subject to editing for clarity and space. Letters submitted, published, or posted elsewhere, in print or online, will be disqualified. To submit a Letter, go to www.submit2science.org.

References

1. J. Haffer, *Science* **165**, 131 (1969).
2. V. Rull, *J. Biogeogr.* **32**, 921 (2005).
3. V. Rull, *Mol. Ecol.* **17**, 2722 (2008).
4. M. Foote, D. M. Raup, *Paleobiology* **22**, 121 (1996).
5. R. T. Pennington *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 13783 (2010).
6. S. Nee *et al.*, *Philos. Trans. R. Soc. London B* **344**, 77 (1994).
7. D. L. Rabosky, *Evolution* **64**, 1816 (2010).
8. N. Cusimano, S. S. Renner, *Syst. Biol.* **59**, 458 (2010).
9. G. E. Shephard *et al.*, *Nat. Geosci.* **3**, 870 (2010).

CORRECTIONS AND CLARIFICATIONS

Perspectives: "The feeding habits of ammonites" by K. Tanabe (7 January, p. 37). The legend should read as follows with corrected genus names: "(**Top**) *Polyptychoceras sp.* with Baculites-like lower and upper jaws… (**Bottom**) *Anagaudryceras limatum* (a lytoceratid) with a nautilus-like lower jaw."

Policy Forum: "Boosting CITES" by J. Phelps *et al.* (24 December 2010, p. 1752). The heading for the fourth potential solution was missing. "A Peer-Review Process" should have appeared before the paragraph on the second page that begins, "CITES shortcomings may be overlooked because the convention lacks internal and external checks and balances." The header has been added in the HTML version online.

News Focus: "Will homebody researchers turn Japan into a scientific backwater?" by D. Normile (10 December 2010, p. 1475). There is a shift in the increments on the vertical axis of the graph indicating the number of individuals making overseas visits. From 0 up to 10,000 individuals, the graph uses increments of 2000; above 20,000, it uses increments of 20,000.

News of the Week: "U.N. biodiversity summit yields welcome and unexpected progress" by D. Normile (5 November 2010, p. 742). The name of Alison Stattersfield, head of science for BirdLife International, was misspelled.

Published by AAAS

Proyectos de investigación-conservación

Resolviendo una incógnita biogeográfica, el caso de la Rand Flora afro-mediterránea.

Mario Mairal[1,2], Andrea Sánchez-Meseguer[1]

[1] Departamento de Biodiversidad y Conservación, Real Jardín Botánico-CSIC, Madrid.

[2] Correspondencia: Mario Mairal (mmairal@rjb.csic.es)

Plaza de Murillo 2, 28014-Madrid, SPAIN. Phone: +34 91 4203017; Fax: +34 914200157.

RESUMEN

Existe un enigmático patrón florístico que ha intrigado a científicos y naturalistas desde hace décadas. Consiste en una distribución de linajes de plantas emparentados que se encuentran distribuidos en regiones florísticas alrededor del continente africano, y separados por áreas de clima hostil. En biogeografía este tipo de distribuciones fragmentadas se conocen como disyunciones.

Este patrón de distribución florística en forma de anillo, se conoce como Rand Flora.

Dos hipótesis se han postulado para explicar este patrón: 1) Una serie de acontecimientos climáticos y geológicos habrían extirpado una flora ancestral de parte de su área de distribución, quedando así relegada a refugios alrededor de África. 2) Las disyunciones observadas son el resultado de recientes eventos de dispersión a larga distancia, con una posterior diversificación en las nuevas áreas.

El objetivo del proyecto es entender los factores históricos que han configurado este patrón florístico. Se aborda el análisis a partir del estudio comparado de varios grupos de plantas que muestran dicha distribución, usando técnicas filogenéticas moleculares, datación y nuevos métodos biogeográficos.

En el presente trabajo se hace una introducción al patrón conocido como Rand Flora. Se enumeran sugerentes ejemplos para interpretar las distribuciones geográficas. Se presentan los estudios de dos casos de especial relevancia, relacionados con dicho patrón; el caso Canarina y el caso Hypericum. Además se exponen y discuten los primeros resultados del proyecto.

Palabras clave: biogeografía, *Canarina*, dispersión, *Hypericum*, Rand flora, vicarianza.

INTRODUCCIÓN

Existe una distribución enigmática de especies en los márgenes del continente africano, conocida como patrón Rand Flora. Dicho patrón (Figura 1) relaciona las floras de regiones tan distantes como Macaronesia (Azores, Madeira, Canarias y Cabo Verde), sur de Arabia, este de África y Sudáfrica (Le Brun 1971; Bramwell 1985; Andrus *et al.* 2004; Sanmartín *et al.* 2010). Se ha propuesto que varios linajes de plantas muestran este patrón de distribución en base a estudios morfológicos y análisis filogenéticos moleculares. En la Tabla 1 se muestra una relación de grupos que ejemplifican esta distribución, aunque podríamos enumerar numerosos linajes más con distribuciones continentales este-oeste (*Parolinia*, *Tolpis*, *Chrysoprenanthes*, *Euphorbia*) u otras plantas macaronésicas con parientes en Sudáfrica (*Justicia*, *Phyllis*,…).

Proyectos de investigación-conservación

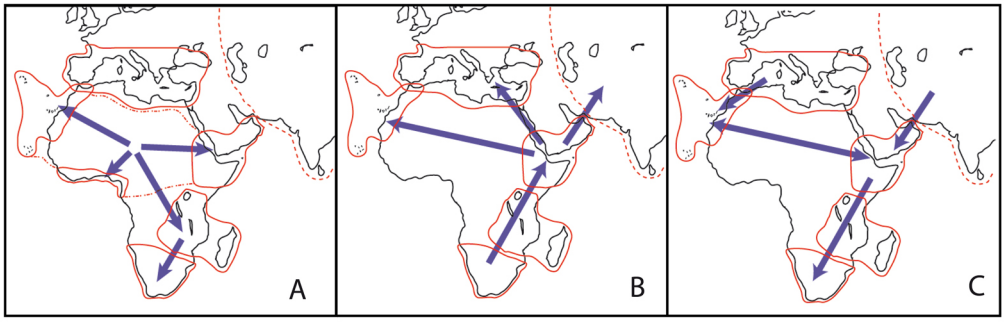
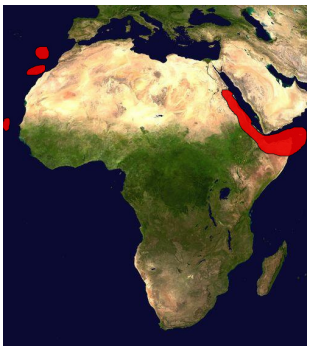
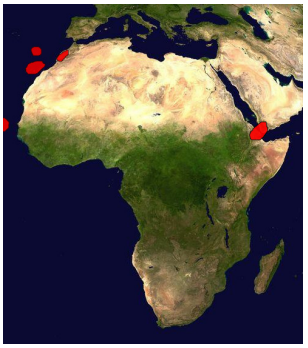
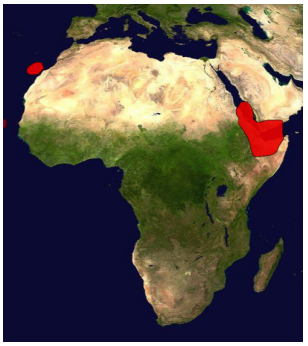
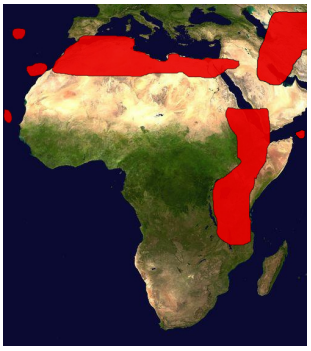
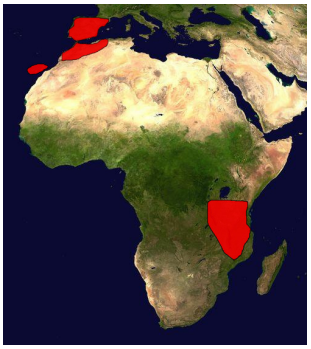
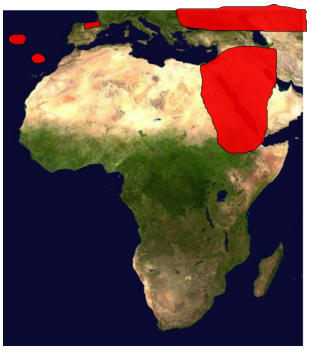
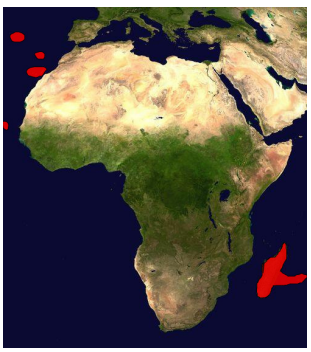
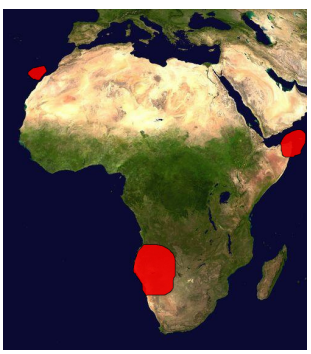


Figura 1. Hipótesis más importantes para explicar el patrón de distribución Rand Flora. A) Vicarianza – los cambios climáticos han producido la fragmentación de una flora continua en el norte de África. B) Dispersión hacia el norte – el patrón se formó por la inmigración de linajes desde el sur de África hacia el norte, y desde ahí hacia el oeste. C) Dispersión hacia el sur – los linajes se dispersaron desde el Mediterráneo y desde Asia hacia el sur a través del este de África. Figura modificada de Sanmartin *et al.* (2010).

Tabla 1. En esta tabla se enumeran y describen algunos linajes que presentan el patrón de distribución Afro-mediterránea conocido como la Rand Flora. Todas las fotos y figuras de la tabla realizadas por M. Mairal Pisa.

 *Dracaena tamaranae* (Canarias)	***Dracaena*** (Asparagaceae). Los legendarios dragos fueron uno de los primeros ejemplos conocidos del patrón. Mientras que la especie *Dracaena draco* aparece en Macaronesia y oeste de África, habría que viajar hasta el este de África o a la isla de Socotra para encontrar cuatro especies más del género (Mwachala, 2005).	
 *Aeonium gorgoneum*. Cabo Verde.	***Aeonium*** (Crassulaceae). Los Bejeques o veroles del género *Aeonium* aparecen en su mayoría distribuidos por la región macaronésica, pero también hay algunas especies en el Éste de África. Son un buen ejemplo de radiación adaptativa en Canarias, con alrededor de 38 especies, 2 especies más en Madeira, 1 en Cabo Verde y 1 en Marruecos. Sólo dos especies del género aparecen en el Este de África (Mort *et al.*, 2002).	
 *Campylanthus glaber* (Cabo Verde)	***Campylanthus*** (Plantaginaceae). El romero marino (*Campylanthus salsoloides*) endémico de Canarias, y *Campylanthus glaber* endémico de Cabo Verde poseen al menos una docena de representantes del mismo género en el Cuerno de África y Oeste de Pakistán (Thiv *et al.*, 2010).	

 *Campanula jacobaea* (Cabo Verde)	***Campanula*** (Campanulaceae). En una de las tribus de las campanillas del género Campanula se observa una disyunción entre islas. *Campanula jacobea* aparece en el archipiélago de Cabo Verde, mientras que 7000 km al este, en la isla de Socotra, aparece Campanula balfouri (Roquet *et al.*, 2009).	
 *Adenocarpus foliolosus* (Canarias)	***Adenocarpus*** (Fabaceae). Un grupo monofilético compuesto por las 3 especies canarias (*A. foliolosus, A. ombriosus* y *A. viscosus*) y una mediterránea (A.complicatus), estaría relacionado con la única especie del género que aparece en el Este de África (*Adenocarpus manii*) (Percy y Cronck, 2002).	
 *Geranium reuteri* (Canarias)	***Geranium subgenus Robertium*** (Geraniaceae). Un grupo de especies de Geranium endémicas de Macaronesia, Marruecos y Península Ibérica componen el grupo hermano de un clado formado por especies del este de África (Fiz *et al.*, 2008).	
 *Sideroxylon canariense* (Canarias)	***Sideroxylon*** (Sapotaceae). Las tres especies macaronésicas, *Sideroxylon mirmulans* (Azores), *Sideroxylon canariense* (Canarias) y *Sideroxylon marginata* (Cabo Verde) forman un clado monofilético con especies de Madagascar e Islas Mascareñas (Smedmarck *et al.*, 2006).	
 *Camptoloma canariensis* (Canarias)	***Camptoloma*** (Scrophulariaceae). Éste género esta formado solamente por tres especies, una de ellas endémica de la isla de Gran Canaria (*Camptoloma canariensis*), otra especie presente en Somalia e Isla de Socotra (*Camptoloma villosa*), y por último otra en Angola y Namibia (*Camptoloma rotundifolia*).	

Proyectos de investigación-conservación

Pero no todas las disyunciones son tan acentuadas, todavía podemos observar distribuciones de especies cuya fragmentación nos da pistas para inferir una biota predesértica. Este sería el caso de algunas especies todavía presentes en puntos intermedios entre el este y el oeste de África, y que aparecen en islas continentales en medio del desierto del Sáhara, como las montañas del Tibesti (Chad) o el macizo del Hoggar (Argelia). He aquí el caso del brezo (*Erica arborea*), la tabaiba dulce (*Euphorbia balsamifera*), o algunas especies de campanillas del género *Campanula*. La presencia de fósiles de algunos de estos taxones en localidades neógenas en el Sáhara podría sugerir que formaban parte de una biota predesértica. Además, hasta hace aproximadamente 4000 años las formaciones de estepa y sabana africanas alcanzaron latitudes mucho más al norte que en la actualidad (Kröpelin 2008; Renssen 2006), permitiendo la existencia de hábitats apropiados en otros lugares del Sáhara y facilitando el flujo de semillas entre estos refugios.

Se han barajado dos posibles hipótesis para explicar el origen de la Rand flora (Figura 1):

1. La hipótesis de vicarianza. Sugiere que las especies actuales son el testigo presencial de una macroflora continental bastamente distribuida en el pasado. Ésta flora habría ocupado una gran parte de la superficie del continente africano, habiendo quedado refugiada en los márgenes del continente como resultado de diversas crisis de aridez desde el Mioceno (23 millones de años; abreviado, Ma) en adelante. Los márgenes del continente habrían actuado como refugios debido a su mayor bonanza climática, en comparación con las fuertes fluctuaciones climáticas acaecidas en el interior del continente. Estos refugios corresponden a los márgenes continentales del este y el oeste de África (Figura 1A; Axelrod & Raven, 1978; Bramwell, 1985; Quezel, 1979, Andrus *et al.* 2004; Thiv *et al.* 2010).

2. La hipótesis de dispersión. Sugiere que la distribución actual es el resultado de eventos recientes de dispersión a larga distancia entre áreas geográficamente aisladas, seguidos de una diversificación de los taxones *in situ*. Para ésta teoría se han propuesto varias rutas de migración; hacia el sur, desde la región mediterránea o desde el oeste de Asia (Figura 1-B1), y hacia el norte, desde el sur de África, vía este de África (Figura 1-B2; Galley *et al*., 2007). Esta hipótesis predice que los linajes estudiados no van a presentar una historia común a lo largo de su historia biogeográfica.

Hasta ahora los estudios biogeográficos sobre la Rand Flora se han limitado a la comparación de los patrones biogeográficos de distintos linajes sin que haya habido un intento de contrastar estas hipótesis estadísticamente. El desarrollo de nuevas herramientas analíticas en inferencia biogeográfica y datación filogenética, junto con la acumulación de datos paleoclimáticos y paleontológicos sobre la vegetación del norte de África, nos permiten por primera vez examinar los factores responsables de esta disyunción continental dentro de un riguroso marco evolutivo.

El presente proyecto de investigación titulado *"Reconstrucción del origen de la Rand Flora Afro-Mediterránea con datación filogenético e inferencia biogeográfica"* tiene como objetivo final inferir los factores históricos que han configurado este enigmático patrón florístico: vicarianza y extinción gradual de una antigua flora terciaria o dispersión y diversificación reciente entre áreas aisladas geográficamente.

Para ilustrar la forma de proceder en el proyecto, en el siguiente apartado se analizan de forma detallada dos de los grupos que se ha propuesto que presentan la distribución "Rand Flora", los géneros *Canarina L.* (Familia Campanulaceae) e *Hypericum L.* (Familia Hypericaceae). Cada uno de estos subproyectos se encuentra en diferentes fases de desarrollo. Finalmente, se presentan los primeros resultados obtenidos en el proyecto, que aún está en curso. Estos resultados son el producto del primer meta-análisis realizado con el patrón biogeográfico de la Rand Flora.

Proyectos de investigación-conservación

EL SUBPROYECTO *CANARINA*

El género *Canarina* se sitúa dentro de la familia de las campanuláceas, pertenece a la tribu *Platycodoneae* (mediante tribu se hace referencia a la categoría taxonómica por debajo de familia) y sólamente comprende 3 especies, que son un modelo manifiesto de disyunción afro-macaronésica; *Canarina canariensis*, aparece en las Islas Canarias occidentales, mientras que para encontrar el resto de especies tenemos que atravesar el desierto del Sáhara hasta el este de África, donde podemos encontrar las otras dos especies: *Canarina abyssinica* y *Canarina eminii* (Figura 2). Esta distribución nos puede servir como organismo modelo para estudiar la disyunción afro-macaronésica. El resto de géneros de la tribu Platycodoneae; Platycodon (1 especie sp), *Codonopsis* (43–55 sp) y *Cyananthus* (44 sp), se distribuyen todas en el este de Asia, convirtiendo a *Canarina* en el único taxón de las Platycodoneae representado en África. Esto, junto a su llamativa disyunción nos plantea una sugerente incógnita biogeográfica.

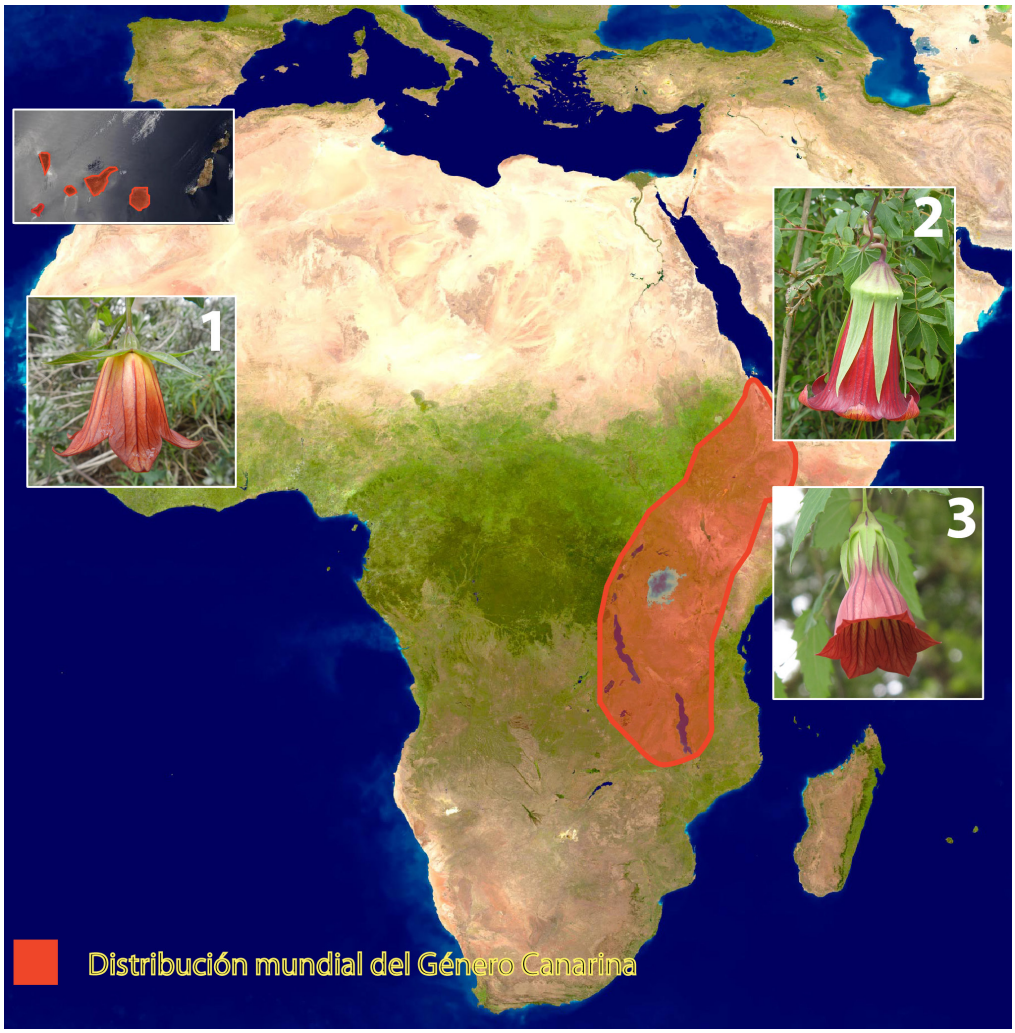


Figura 2. Mapa con la distribución actual de las especies del género *Canarina*.

1. *Canarina canariensis*, presente en las islas Canarias occidentales.

2 y 3. C*anarina eminii* y *Canarina abyssinica*, con distribución solapada en el este de África.

Figura realizada por M. Mairal Pisa.

Proyectos de investigación-conservación

Para comprender esta distribución es importante conocer el marco paleoclimático de África. Sabemos que hace doce millones de años, en el Mioceno medio, el actual desierto del Sáhara estaba cubierto por sábana tropical (Quezel, 1978), pudiéndose encontrar sábanas arboladas hasta hace solamente 5000 años (Kröpelin 2008; Renssen 2006). De esta manera, las actuales biotas (conjunto de especies que ocupan un determinado área) macaronésicas y del este de África podrían haber estado conectadas. Es probable que esta conexión haya permitido que los frutos carnosos de Canarina, los cuales son consumidos por aves, hayan sido dispersados a Canarias por ornitocoria (dispersión de semillas realizada por pájaros). Posteriormente *Canarina* se habría extinguido del norte de África cuando estos bosques húmedos desaparecieron por la creciente aridificación del territorio. A falta de registro fósil, a priori no podemos saber si *Canarina* vivió a lo largo de las extensas sabanas africanas o se produjo un evento de dispersión a larga distancia.

Dentro de este contexto se pretende dilucidar la historia evolutiva de este género, para lo que se está haciendo un importante esfuerzo de muestreo y trabajo molecular. En concreto, estamos trabajando con tres métodos filogenéticos moleculares diferentes: análisis de regiones del genoma nuclear y plastidial, análisis de microsatélites y AFLPs (Polimorfismos en la Longitud de Fragmentos Amplificados). Estas técnicas permiten conocer las relaciones de parentesco entre las especies de *Canarina* en base a las distancias genéticas que existen entre ellas, así como el flujo génico que se ha producido entre los individuos de sus poblaciones. En último término, toda esta información nos permitirá inferir el área de origen del grupo, la dirección de la dispersión y la antigüedad de la separación entre las poblaciones a ambos lados del continente africano.

Las dos especies de *Canarina* del este de África se pueden encontrar en las selvas afromontanas húmedas, hábitat en un alarmante estado de retroceso, que ha quedado reducido a retazos de selvas dispersa por el este de África. Mediante el estudio de secuencias de ADN en relación a la distribución geográfica de los individuos se pueden también mejorar los esfuerzos de conservación, identificando las regiones en las que las especies tienen una diversidad genética más alta o exclusiva, y que por tanto deberían ser priorizadas en las políticas de conservación.

Pero además, durante el curso de esta investigación, han descollado otras preguntas interesantes relacionadas con el objeto de estudio. Sabemos que las especies de *Canarina* son polinizadas por pájaros (Olessen, *et al.*, in press), de esta manera, poseen diversos caracteres relacionados con la polinización ornitófila; flores robustas, néctar más diluido, colores que no reflejan la luz ultravioleta. Su interacción con los polinizadores es compleja; mientras que en África, *Canarina* es polinizada por pájaros de la familia Nectariniidae especializados en consumir polen (por ejemplo *Cinnyris mariquensis*), en Canarias, sin embargo, es polinizada por pájaros de dieta generalista, como el mosquitero canario (*Phylloscopus canariensis*) o el herrerillo (*Cyanistes caeruleus*). La ausencia de especies de pájaros con dieta especialista en Macaronesia nos sugiere fascinantes preguntas evolutivas, por ejemplo ¿por qué poseen adaptaciones a polinizadores especialistas estas plantas insulares? Así, pretendemos comparar por primera vez la biología reproductiva y los efectos de la polinización ornitófila entre isla y continente. Con ésto se pretende dilucidar las interacciones ecológicas entre estos organismos, y la importancia que podría tener la extinción de alguna de sus funciones en el ecosistema, ya de por si muy fragmentado.

Proyectos de investigación-conservación

EL SUBPROYECTO *HYPERICUM*

Hypericum L. es otro de los géneros que se ha propuesto que muestran el patrón de distribución Rand Flora. El grupo contiene más de 470 especies distribuidas en todos los continentes y diferentes ecosistemas, y solamente falta en los polos, en los desiertos y en zonas bajas tropicales (Robson, 1977; Figura 3). En África, *Hypericum* está presente en la mayor parte de los hábitats excepto desiertos y zonas bajas tropicales. *Hypericum* es un grupo muy interesante para nuestro estudio porque se trata de un género muy antiguo. Tanto el registro fósil como las dataciones moleculares, indican que el orden Malpighiales al que *Hypericum* pertenece, divergió rápidamente a partir del Cretácico Medio (110 Ma) (Davis *et al.*, 2005), y datan la separación de *Hypericum* de su grupo hermano Vismia en el Cretácico-Paleoceno, hace 65 Ma. En una revisión reciente del registro fósil de *Hypericum*, Sánchez-Meseguer y Sanmartin (2012) sostienen que el origen del género se remonta al Eoceno superior. Entre otros restos fósiles, se ha encontrado polen fósil de *Hypericum* del Oligoceno inferior en España, aprox. 30 Ma. Esta información nos indica que tanto el género, como su presencia en África son probablemente anteriores al supuesto origen Mioceno de la "Rand Flora" y es por ello un grupo idóneo para nuestro estudio.

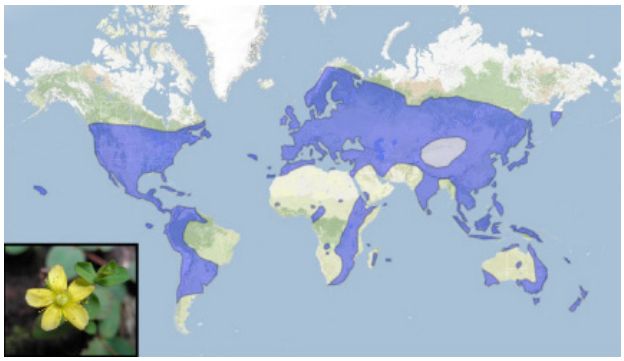


Figura 3. Mapa con la distribución actual de las especies del género *Hypericum* modificado de Robson (1977). Figura realizada por A. Sanchez-Meseguer.

Robson (1981) describió varias disyunciones afro-mediterráneas para *Hypericum* basándose en el estudio de caracteres morfológicos. Entre otras, propuso que *Hypericum* roeperianum (distribuido en el este África) está emparentado con *H. canariense* (Macaronesia), o que H. quartinianum (este de África) está relacionado con *H. glandulosum* (Macaronesia).

Al igual que ocurre con otros grupos incluidos en la Rand Flora, la existencia de disyunciones ha sido propuesta en base a caracteres morfológicos, pero no existe una filogenia molecular que nos permita testar estas hipótesis de parentesco en base a caracteres moleculares. Además, los estudios previos filogenéticos del género *Hypericum* eran muy reducidos en número de especies y estaban centrados en áreas geográficas concretas como el este de Asia y norte de América (Crockett *et al.*, 2004; Park y Kim, 2004; Pilepic *et al.*, 2011). Por ello, para este proyecto hemos generado la primera filogenia molecular representativa para todo el género que contiene representantes de todos los lugares donde *Hypericum* está distribuido (Sánchez-Meseguer *et al.*, en revisión) y de toda la variación morfológica que contiene el grupo. Además, hemos estimado los tiempos desde que divergieron los linajes de *Hypericum* (Sánchez-Meseguer *et al.*, en preparación). Los resultados obtenidos aún no han sido publicados, pero de forma preliminar se puede concluir que algunas de las especies que se ha propuesto que muestran el patrón de distribución afro-mediterráneo no están emparentadas entre si (Sánchez-Meseguer *et al.*, en revisión). Éste es el caso para la supuesta relación entre *H. canariense* e *H. roeperianum*. La hipótesis de disyunción que postulaba Robson no se confirma, ya que ambas especies aparecen agrupadas en clados distintos (no son especies relacionadas). H. canariense está emparentada con taxones mediterráneos, mientras que H. roeperianum desciende de linajes distribuidos en Asia, y por lo tanto su presencia en el este de África probablemente se debe a una dispersión desde Asia. Lo mismo ocurre entre *H. quartinianum* y *H. glandulosum*. Por lo tanto, la distribución de estos linajes en los márgenes del continente africano no podría ser explicada por la hipótesis de fragmentación de una antigua flora continental Africana.

Sin embargo, las evidencias moleculares sugieren que sí que existe relación entre *Hypericum tomentosum* (distribuido en el norte de África y en la Península Ibérica) e *Hypericum somaliense* (Somalia). Ambas especies aparecen emparentadas dentro de un clado mayormente mediterráneo. La divergencia entre ellas data del Plioceno (Sánchez-Meseguer *et al*., en preparación), lo que nos indica que pudieron formar parte de la Rand Flora, aunque la edad de separación entre ellas fue más reciente que en otros casos estudiados (Sánchez-Meseguer *et al*., en preparación).

PRIMEROS RESULTADOS DEL PROYECTO

En 2010 Sanmartín y colaboradores presentaron el primer meta-análisis del patrón Rand Flora. En él se incluían 13 linajes de plantas con distribución disyunta en los márgenes del continente africano. Los linajes incluidos fueron: *Aeonium* (Crassulaceae), *Adenocarpus* (Fabaceae), *Androcymbium* (Colchicum) (Colchicaceae), *Convolvulus* (Convolvulaceae), *Monsonia* (Geraniaceae), *Moraea* (Iridaceae), *Sideroxylon* (Sapotaceae), *Ceropegieae* (Apocynaceae), *Geranium* (subgenero Robertium) (Geraniaceae), *Solanum* (subgenero *Leptostemonum* p.p.) (Solanaceae), dos grupos del gran género *Euphorbia* (subg. *Rhizanthium* y subg. *Esula* p.p) (Euphorbiaceae), así como la familia Campanulaceae.

En el artículo se explora el uso del método de Biogeografía Bayesiana de Islas (BIB) propuesto por Sanmartín *et al.* (2008) para estimar la tasa de intercambio biótico o flujo génico entre áreas geográficamente aisladas. Este método usa datos sobre la distribución geográfica de las especies y datos moleculares de múltiples grupos taxonómicos, que difieren en su edad de origen, tasa de evolución molecular y capacidad dispersiva.

En el trabajo de Sanmartín *et al.* (2010) se pudo establecer que las tasas de intercambio biótico resultaron ser más altas entre el este y el oeste de África que por ejemplo entre norte-este o norte-sur de África. Ésto sugiere que las áreas del este y el oeste permanecieron más tiempo conectadas, o dicho de otra manera, que el intercambio biótico entre estas dos áreas duró hasta tiempos más recientes que entre las otras áreas estudiadas. El intercambio biótico entre Sudáfrica con el este o el norte de África fue probablemente más antiguo. Cuanto más antiguo es un intercambio biótico entre dos áreas es más probable que la extinción haya eliminado su rastro, siendo más difícil observar su señal en una filogenia. Sin embargo, si el intercambio es más reciente, la extinción todavía no ha tenido tiempo de eliminar esa señal.

Además, estos datos concuerdan con la información geológica. Sabemos que la desertificación del norte de África fue posterior a la del sur y el este de África, lo que esclarece una pieza más de este complejo puzzle evolutivo.

La alta tasa de dispersión detectada entre el oeste (Macaronesia-noroeste de África) y este de África (Cuerno de África-Península Arábiga) apoyaría la hipótesis de vicarianza, dónde una flora macrocontinental quedaría fragmentada por las variaciones climáticas ocurridas desde el Mioceno (Axelrod & Raven, 1978), dándose a posteriori una diversificación in situ. Algunos ejemplos cuya distribución se explica por un fenómeno de vicarianza son los géneros: *Aeonium*, *Campylanthus* o *Adenocarpus* (Tabla 1).

Sin embargo, sería improbable pensar que todo ha sucedido por un único evento de vicarianza, estas tasas reflejarían eventos repetidos de dispersión y vicarianza coincidiendo con la alternancia de ciclos húmedos y áridos en el norte de África desde el Mioceno (Thiv *et al*., 2010), los cuáles habrían permitido el intercambio génico de forma intermitente.

Asimismo, se encontró que la región Macaronesia-noroeste de África presenta una capacidad de carga menor que otras regiones; esto, unido a su mayor tasa de dispersión, sugiere

que la flora Macaronésica está compuesta por linajes que llegaron por migración a larga distancia. La migración de linajes se produjo principalmente desde el Mediterráneo, y en menor grado desde la Península Arábiga. La alta capacidad de carga observada en Sudáfrica junto con la menor tasa de dispersión, indica que la rica diversidad sudafricana surgió en dicho lugar, lo que es respaldado por la estabilidad climática de este área desde el Mioceno (Linder, 2005), algunos géneros representativos serían *Androcymbium*, *Monsonia* o *Moraea*. Ésto señala la región Sudafricana como fuente de linajes, que migrarían hacia el norte vía este de África, subiendo por las Montañas del Drakensberg y el Gran Rift, cuya formación también se daría en el Plioceno (20 millones de años). De esta manera, parte de la rica diversidad afromontana del este de África posiblemente provenga de la región de El Cabo sudafricana (Linder, 2005; Galley *et al*., 2007). Los resultados indican también que ha habido poco intercambio biótico entre las regiones del sur de África y del norte de África, ocurriendo éste por el este de África.

El trabajo presentado por Sanmartin *et al.* (2010) se está continuando en la actualidad, incorporando en el análisis otros muchos grupos de estudio con distribución Rand Flora, entre ellos lo género descritos *Canarina* e *Hypericum*. Se prevé que al término del proyecto de investigación, se tendrá una idea más realista de los factores históricos que han contribuido a la formación del patrón florístico de la Rand Flora.

Los resultados de este proyecto pueden ayudarnos a entender el papel de la extinción asociada a cambios climáticos en la formación de patrones de diversidad vegetal. Asimismo, pueden ser de utilidad en el diseño de políticas de conservación en estas áreas, a través del desarrollo de modelos predictivos para conocer las consecuencias de la *aridificación* actual en el Norte de África y Sur de Europa en la evolución futura de su diversidad vegetal.

AGRADECIMIENTOS

Este trabajo ha sido realizado gracias a la financiación del Ministerio de Ciencia y Educación (proyecto CGL2009-13322-C03-01/BOS) y a las becas de doctorado de A.S.M (AP-2007-01698) y a M.M (BES-2009-028958). Los autores quieren agradecer la ayuda prestada para la preparación del manuscrito a los responsables del proyecto, I. Sanmartín y J.J. Aldasoro, y también a M. Alarcón y C.L. Anderson. Además, agradecemos la ayuda técnica a I. Marques, E. Cano, F. Durán, G. Andreu y G. Sanjuanbenito.

BIBLIOGRAFÍA

– ANDRUS, N., TRUSTY, J., SANTOS-GUERRA, A., JANSEN, R. K., FRANCISCO-ORTEGA, J. 2004. Using molecular phylogenies to test phytogeographical links between East/South Africa, Southern Arabia and the Macaronesian islands— a review, and the case of *Vierea* and *Pulicaria* sect. *Vieraeopsis. Taxon*, 53: 333–346.

– AXELROD, D. I., RAVEN, P. H., 1978. Late Cretaceous and Tertiary vegetation history of Africa. *Biogeography and ecology of Southern Africa* (ed. M. J. A. Werger), pp. 77–130. The Hague, The Netherlands: Junk.

– BRAMWELL, D., 1985. Contribución a la biogeografía de las Islas Canarias. Bot. *Macaronésica*, 14: 3–34.

Proyectos de investigación-conservación

– CROCKETT, S.L., DOUGLAS, A.W., SCHEFFLER, B.E., KHAN, I.A., 2004. Genetic profiling of *Hypericum* (St. John's wort) species by nuclear ribosomal ITS sequence analysis. *Pl. Med.*, 70: 1–7.

– DAVIS, C.C., WEBB, C.O., WURDACK, K.J., JARAMILLO, C.A., DONOGHUE, M.J., 2005. Explosive radiation of Malpighiales supports a mid-Cretaceous origin of modern tropical rain forests. *Amer. Naturalist*, 165: E36–E65.

– FIZ, O., VARGAS, P., ALARCÓN, M., AEDO, C., GARCÍA, J. L. Y ALDASORO, J. J., 2008. Phylogeny and historical biogeography of Geraniaceae in relation to climate changes and pollination ecology. *Systematic Botany*, 32.

– GALLEY, C., BYTEBIER, B., BELLSTEDT, D. U., LINDER, H. P., 2007. The Cape element in the Afrotemperate flora: from Cape to Cairo?. *Proc. R. Soc.*, B 274: 535–543.

– KRÖPELIN, S., VERSCHUREN, D., LÉZINE, A.M., EGGERMONT,H., COCQUYT, C., FRANCUS,P., CAZET, J.P., FAGOT,M., RUMES,B., RUSSELL, J.M., DARIUS,F., CONLEY, D.J., SCHUSTER,M., VON SUCHODOLETZ, H., ENGSTRÖM, D.R., 2008. Climate-driven ecosystem succession in the Sahara: The past 6000 years. *Science* 320 : 765–768.

– LE BRUN, J. P., 1971. Quelques phanerogames africaines a` aire disjointe. *Mitteil. Bot. Staatssamm. Mün.*, 10: 438–448.

– LINDER, H. P., 2005. Evolution of diversity: the Cape Flora. *Trends Plant. Sci.*, 10: 536–541.

– Molero, J., Garnatje, T., Rovira, A., Garcia-Jacas, N., Susanna, A., 2002. Karyological evolution and molecular phylogeny in Macaronesian dendroid spurges (*Euphorbia subsect Pachycladae*). *Plant. Syst. Evol.*, 231: 109-132.

– MORT, M. E, SOLTIS D. E., SOLTIS P. S., FRANCISCO-ORTEGA, J. Y SANTOS GUERRA, A., 2002. Phylogenetics and evolution of the Macaronesian Crassulaceae inferred from nuclear and chloroplastic sequences. *Systematic Botany*, 27: 271-288.

– MWACHALA, G. 2005. Systematics and ecology of *Dracaena* L. (Ruscaceae) in central, east and southern Africa. PhD Dissertation, Univeristy of Koblenz-Landau, Institute of Biology.

– OLESEN, J.M., ALARCÓN, M., EHLERS, B.K., ALDASORO, J.J., AND ROQUET, C. 2012. Pollination, biogeography and phylogeny of oceanic Island bellflowers (Campanulaceae). Perspectives in *Plant Ecology, Evolution and Systematics*, in press.

– PERCY, D. M., CRONK, Q. C. B., 2002. Different fate of island brooms; contrasting evolution in *Adenocarpus*, Genista and Teline in the Canary Islands and Madeira. *Am. J. Bot.* 89: 854-864.

– PARK, S.J., KIM, K.J., 2004. Molecular phylogeny of the genus Hypericum (Hypericaceae) from Korea and Japan: Evidence from nuclear rDNA ITS sequence data. *J. Pl. Biol.*, 47: 366–374.

– PILEPIĆ, K.H., BALIĆ, M., BLAŽINA, N., 2011. Molecular phylogenetic relationships of some *Hypericum* (Hypericaceae) based on ITS sequences. *Plant Biosystems*, 145: 81–87.

– Quezel, P., 1979. Analysis of the flora of Mediterranean and Saharan Africa. *Ann. Missouri Bot. Gard.*, 65: 479-534.

– RENSSEN, H., BROVKIN, V., FICHEFET, T., GOOSSE, H., 2006. Simulation of the Holocene climate evolution in Northern Africa: The termination of the African Humid Period. *Quat. Int.*, 150: 95–102.

– ROBSON, N.K.B., 1977. Studies in the genus *Hypericum* L. (Guttiferae). 1. Infrageneric classification. *Bull. Brit. Mus. (Nat. Hist.), Bot.*, 5: 295–355.

– ROBSON, N.K.B., 1981. Studies in the genus *Hypericum* L. (Guttiferae). 2. Characters of the genus. *Bull. Brit. Mus. (Nat. Hist.), Bot.*, 8: 55–226.

– ROQUET, C., SANMARTÍN, I., GARCIA-JACAS, N., SÁEZ, LL., SUSANNA A.,WIKSTRÖM N.,ALDASORO, J.J., 2009. Reconstructing the history of Campanulaceae with a Bayesian approach to molecular dating and dispersal–vicariance analyses. *Molecular Phylogenetics and Evolution*, 52: 575–587.

– SÁNCHEZ-MESEGUER, A., ALDASORO, J.J., SANMARTÍN, I. (en revisión). New plastid and nuclear DNA evidence help solve relationships within the large and cosmopolitan genus *Hypericum* (Hypericaceae).

– SÁNCHEZ-MESEGUER, A., SANMARTÍN, I. Paleobiology of the genus *Hypericum* (Hypericaceae): a survey of the fossil record and some palaeogeographic implications. Anales del Jardín Botánico de Madrid (aceptado, fecha prevista de publicación 2012).

– SANMARTÍN, I., VAN DER MARK, P., RONQUIST, F. 2008. Inferring dispersal: a Bayesian, phylogeny-based approach to island biogeography, with special reference to the Canary Islands. *J. Biogeo.*, 35: 428–449.

– SANMARTÍN, I., ANDERSON, C.L., ALARCON, M., RONQUIST, F., ALDASORO, J.J., 2010. Bayesian island biogeography in a continental setting: the Rand Flora case. *Biol. Lett.* 6, 703–707.

– SMEDMARK, J. E. E., SWENSON, U., ANDERBERG, A., 2006. Accounting for variation and substitution rates through time in Bayesian phylogeny reconstruction of Sapotoideae. *Mol. Phyl. Evol.*, 39: 706-721.

– THIV, M., THULIN, M., HJERTSON, M., KROPF, M., LINDER, H. P. 2010. Evidence for a vicariant origin of Macaronesian- Eritreo/Arabian disjunctions in *Campylanthus* Roth (Plantaginaceae). *Mol. Phyl. Evol.*, 54: 607–616.